

ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2017/18

DOCTORANDO: **MATOS FLORES, RAÚL**
D.N.I./PASAPORTE: ******26801**

PROGRAMA DE DOCTORADO: **D444 TECNOLOGÍAS DE LA INFORMACIÓN GEOGRÁFICA**
DEPARTAMENTO DE: **GEOLOGÍA, GEOGRAFÍA Y MEDIO AMBIENTE**
TITULACIÓN DE DOCTOR EN: **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

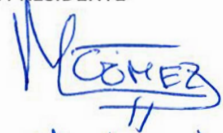
En el día de hoy 03/10/2017, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de MARÍA JESÚS SALADO GARCÍA.

Sobre el siguiente tema: *SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y MODELIZACIÓN DEL VALOR DEL SUELO RESIDENCIAL*

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL¹ de (no apto, aprobado, notable y sobresaliente): **SOBRESALIENTE**

Alcalá de Henares, 3 de octubre de 2017


EL PRESIDENTE


Fdo.: Montserrat Gómez

EL SECRETARIO


Fdo.: Raúl Romero

EL VOCAL


Fdo.: José Luis García

Con fecha 31 de octubre de 2017 la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- ☐ Conceder la Mención de "Cum Laude"
☒ No conceder la Mención de "Cum Laude"

FIRMA DEL ALUMNO,

La Secretaria de la Comisión Delegada




Fdo.:

¹ La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:



Universidad
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 31 de octubre, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por *MATOS FLORES, RAÚL*, el día 3 de octubre de 2017, titulada *SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y MODELIZACIÓN DEL VALOR DEL SUELO RESIDENCIAL*, para determinar si a la misma se le concede la mención "cum laude", arrojando como resultado, 3 votos en contra.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve no otorgar la Mención de "cum laude"** a dicha Tesis.

Alcalá de Henares, 2 de noviembre de 2017

EL PRESIDENTE DE LA COMISIÓN DE ESTUDIOS
OFICIALES DE POSGRADO Y DOCTORADO



Juan Ramón Velasco Pérez

Copia por e-mail a:

Doctorando: MATOS FLORES, RAÚL

Secretario del Tribunal: RAÚL ROMERO CALCERRADA.

Directora de Tesis: MARÍA JESÚS SALADO GARCÍA



Universidad
de Alcalá

ESCUELA DE DOCTORADO
Servicio de Estudios Oficiales de
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D^a _____
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo
presentado la misma en formato: ☐ soporte electrónico ☐ impreso en papel, para el depósito de la
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: _____ se procede, con
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a _____ de _____ de 20____



Fdo. El Funcionario



Universidad
de Alcalá

DEPARTAMENTO DE GEOLOGÍA, GEOGRAFÍA Y
MEDIO AMBIENTE

C/ Colegios, 2
28801 Alcalá de Henares (Madrid)
Teléfonos: 91 885 44 29
Fax: 91 885 44 39
e-mail: dep419@uah.es

Francisco Javier Salas Rey, Director del Departamento de Geología, Geografía y Medio Ambiente de la Universidad de Alcalá, en representación de su Consejo de Departamento,

HACE CONSTAR

Que la Tesis Doctoral titulada “Sistemas de Información Geográfica y Modelización del Valor del Suelo Residencial”, realizada por D. **Raúl Matos Flores**, reúne las características adecuadas para su exposición y defensa pública: rigor científico, aportaciones novedosas y aplicación adecuada de la metodología. Por tales motivos doy mi conformidad para que se proceda al depósito y registro de esta Tesis Doctoral.

Alcalá de Henares, 7 de abril de 2017



Francisco Javier Salas Rey
Director del Departamento de Geología, Geografía y Medio Ambiente



Universidad
de Alcalá

DEPARTAMENTO DE GEOLOGÍA, GEOGRAFÍA Y
MEDIO AMBIENTE

C/ Colegios, 2
28801 Alcalá de Henares (Madrid)
Teléfonos: 91 885 44 29
Fax: 91 885 44 39
e-mail: dep419@uah.es

María Jesús Salado García, Profesora Titular de Geografía Humana, como directora de la tesis Doctoral realizada por D. Raúl Matos Flores y titulada "Sistemas de Información Geográfica y Modelización del Valor del Suelo Residencial",

Informa de que dicha tesis reúne los requisitos propios de este tipo de trabajo: rigor científico, aportaciones novedosas y aplicación adecuada de la metodología. Por tal motivo da el VISTO BUENO para su presentación.

Alcalá de Henares, 6 de abril de 2017

María Jesús Salado García
Directora de la investigación





Programa de Doctorado en Cartografía, Sistemas de Información Geográfica y
Teledetección

Sistemas de Información Geográfica y Modelización del Valor del Suelo Residencial

Tesis Doctoral presentada por:

Raúl Matos Flores

Bajo la dirección de:

Dra. María Jesús Salado García

Alcalá de Henares, Marzo de 2017

A Marisol, Andrés y Laura

Agradecimiento

Llegar al punto de presentar los resultados de la presente investigación en gran medida se debe al apoyo de aquellos que han creído en este esfuerzo. En primer lugar, deseo agradecer a mi directora de tesis, la Dra. María Jesús Salado García, por la confianza y ayuda recibida durante todo el proceso de la investigación. Agradezco inmensamente sus atinadas correcciones y sugerencias realizadas al trabajo, sobre todo, por haberlas realizado comprendiendo mi alejamiento al proceso de investigación en momentos arduos de mi vida.

Al Departamento de Geología, Geografía y Medio Ambiente de la Universidad de Alcalá, a su programa de doctorado en “Cartografía, Sistemas de Información Geográfica y Teledetección” y a su equipo docente. Gracias a ustedes perfeccioné mi formación en las Tecnologías de la Información Geográfica y mi capacidad investigadora.

A todos mis compañeros del programa de Doctorado, especialmente y con mucho afecto a Luis Marino Santana y Alexander Ariza. Más que por lo académico, muchas gracias por su amistad y los momentos compartidos. Nuestras conversaciones existenciales me llevaron a pensar mucho sobre el rol de Puerto Rico en nuestra América. Se han convertido en dos hermanos.

Especial agradecimiento al Dr. Carlos Guilbe del Departamento de Geografía de la Universidad de Puerto Rico y al Dr. Carlos González de la Universidad Politécnica de Puerto Rico. Sus recomendaciones en el tema de investigación y el ánimo a no detenerme y seguir con este esfuerzo siempre estuvieron presentes. Por eso, muchas gracias a ambos.

A mis padres, Lucy y Raúl. Esta meta lograda también es de ustedes. A todos mis hermanos y sus familias por animarme a seguir adelante en todo momento.

Finalmente, a los seres más especiales en mi vida, mi esposa Marisol y mis hijos Andrés y Laura. Gracias a los tres por haber compartido este camino. Han sido muchos los sacrificios sufridos por todos. Ustedes me han dado la fuerza y el ánimo necesario para cumplir con esta meta de vida. No saben lo agradecido que estoy de ustedes.

Y a todos los que de alguna u otra forma han apoyado este esfuerzo o han colaborado con el mismo, muchas gracias.

Índice General

Resumen.....	I
Abstract.....	II
1. Valoración de Inmuebles con Fines Contributivos en Puerto Rico: Una Mirada Histórica ..	4
1.1 Introducción	4
1.2 Evolución de la contribución sobre la propiedad inmueble en Puerto Rico.	5
1.3 El modelo de Puerto Rico comparado.....	10
1.4 El Problema del Modelo Actual	15
1.5 Como intervenir en la Situación Actual: Objetivos de la Investigación	17
1.6 Estructura de la Investigación	19
2. Valoración Inmobiliaria: Conceptos Básicos.....	22
2.1 El suelo como mercancía.....	22
2.2 Fundamentos de la valoración de inmuebles.....	25
2.3 El Proceso de Valoración	26
2.3.1 Enfoques de Valoración de Inmuebles	28
2.3.2 Valoración Hedónica	31
3. Modelación de Valores de Venta de Inmuebles Residenciales en un Entorno SIG.	35
3.1 Valoración Hedónica en un entorno SIG	36
3.2 Mínimos Cuadrados Ordinarios	36
3.3 Regresión Geográficamente Ponderada	45
4. Caguas: Una Mirada a su Desarrollo Urbano	50
4.1 Localización	50
4.2 Evolución Urbana.....	51
4.2.1 De los orígenes a finales del siglo XX	51
4.2.2 Caguas Contemporáneo.....	58
4.3 Tipología edificatoria en Caguas.....	63
5. Análisis Exploratorio de Datos para la Predicción de Valores Inmuebles	73

5.1 Fuente de los datos	75
5.2 Diagnóstico del modelo de regresión inicial	81
5.3 Diagnóstico del modelo de regresión con datos muestreados (Valor Mediano).....	97
5.4 Selección Final de Variables	109
6. Valoración Hedónica con Mínimos Cuadrados Ordinarios y Regresión Geográficamente Ponderada: Resultados de un Estudio Comparado	113
6.1 Construcción del modelo con la técnica de los Mínimos Cuadrados Ordinarios.....	113
6.2 Construcción del modelo con la técnica de Regresión Geográficamente Ponderada ..	121
6.3 Comparación de Modelos.....	140
6.4 Selección de Modelo	142
7. Validación del Modelo.....	144
8. Conclusiones	152
Bibliografía	160
Anexo.....	166
ANALISIS DE LA INFLUENCIA DE COMPONENTES GEOGRAFICOS Y	
TEMPORALES EN VALORES RESIDENCIALES UTILIZANDO SIG. Comunicación en	
la Semana Geomática Internacional 2013, Instituto Geográfico Agustín Codazzi (IGAC).	
Bogotá, Colombia	167

Resumen

El suelo constituye una de las bases sobre la que se apoyan las actividades sociales y económicas de la sociedad. Para la administración pública su valor puede representar un ingreso monetario significativo en términos de recaudos por medio de impuestos. En Puerto Rico la última tasación general de propiedades inmuebles se realizó hace poco más de medio siglo. Los criterios y procesos adoptados para los procesos de valoración de inmuebles datan de 1957, por lo tanto, al tasar nuevos inmuebles se utiliza métodos de aproximación basados en premisas que hace décadas dejaron de ser válidas. Con el objetivo de establecer un método que permitiera modernizar el sistema de valoración inmueble se decidió adoptar una técnica basada en la predicción valores inmuebles utilizando datos de ventas residenciales efectuadas en tiempo reciente. Para esto se compararon las técnicas de valoración hedónica utilizando los Mínimos Cuadrados Ordinarios y la técnica de Regresión Geográficamente Ponderada. Para ello se procesaron datos de ventas residenciales en un período cinco años (2005 al 2009) en un Sistema de Información Geográfica. De los resultados obtenidos se concluyó que la Regresión Geográficamente Ponderada permite realizar predicciones más certeras.

Con el análisis de los modelos también se buscó estudiar el efecto de eventos particulares en los procesos de venta de inmueble. En este caso se tomaron los datos de ventas residenciales entre los años 2005 y 2009 porque se buscaba analizar el efecto que podría tener el estallido de la burbuja inmobiliaria en que se encontraba inmerso el mercado inmobiliario de Puerto Rico en el año 2007.

Además, se buscó estudiar la vigencia a largo plazo de los valores de predicción obtenidos por los modelos. Para esto se procedió a obtener datos de ventas residenciales para los cinco años posteriores a los seleccionados para el análisis de los modelos; esto es, datos de ventas residenciales para la zona de estudio entre los años 2010 y 2014. Se tomaron los datos de predicción de los años 2005 al 2009 y se asumió que estos serían las predicciones de valor inmueble de referencia para los años subsiguientes. Calculando el Coeficiente de Dispersión para cada año se analizaron los resultados y se concluyó que las predicciones mantienen su validez por un período de tres años. Con toda la información obtenida y en comparación con modelos de la región se establecieron los lineamientos generales que requiere la actualización del modelo de valoración de inmuebles con fines contributivos en Puerto Rico apoyándose en el uso de Sistemas de Información Geográfica.

Abstract

Land constitutes one of the bases on which social and economic activities of society are supported. For a public administration its value can represent a significant monetary income in terms of tax revenues. In Puerto Rico the last general appraisal of real estate for tax revenues was made just over half a century ago. The criteria and processes adopted for property valuation processes date from 1957, therefore nowadays, when appraising new properties, methods used are based on premises that are no longer valid. In order to establish a method to modernize the island property valuation system, this research pretends to adopt a technique based on the prediction of residential sales made in recent times. For this, a comparison of hedonic valuation techniques using the Ordinary Least Squares and the Geographically Weighted Regression was made. Using Geographic Information Systems, residential sales data from a period of five years (2005 to 2009) were georeferenced. Both techniques were applied and from the results obtained it was concluded that the Geographically Weighted Regression allows a more accurate prediction of land value data.

The analysis of the models also sought to study the impact of particular events on the sale of real estate. In this case the data of residential sales running from 2005 to 2009 wanted to analyze the effect the burst of the 2007 real estate bubble had in the real estate market of Puerto Rico. In addition, the study wanted to analyze the long-term validity of the prediction values obtained by the models. For this, data of residential sales for the years 2010 thru 2014 was georeferenced. Prediction data for the years 2005 to 2009 were taken and adopted as reference of real estate predictions for subsequent years. Coefficient of Dispersion was calculated for each year. It was concluded that the predictions remain valid for a period of three years. With all the information obtained and in comparison with models of the region, general guidelines required to update the property valuation model for tax purposes in Puerto Rico were established. The proposed system is based on the use of Geographic Information Systems.

Capítulo 1

Valoración de Inmuebles con Fines Contributivos en Puerto Rico: Una Mirada Histórica

1. Valoración de Inmuebles con Fines Contributivos en Puerto Rico: Una Mirada Histórica

1.1 Introducción

El suelo constituye una de las bases sobre la que se apoyan las actividades sociales y económicas de la sociedad. Se considera el recurso básico requerido para el crecimiento urbano y a la vez es la fuente principal para la producción agrícola. Para un sistema gubernamental su valor puede representar un ingreso monetario significativo en términos de recaudos por medio de impuestos. Este recaudo solo puede estar garantizado con la existencia de bases de datos fiables que permitan al sector gubernamental conocer a quién pertenece el suelo y cuál es su valor. Esta fuente debe facilitar la adquisición de información necesaria sobre valores de renta y venta del suelo, los cuales forman la base para la toma de decisiones sobre transacciones que afectan el mercado inmobiliario.

La visión del suelo como elemento de valor económico se remonta a tiempos de los egipcios cuando estos establecieron los sistemas de cobro de impuestos sobre la propiedad. Indica Larsson (1991) que la referencia directa más antigua se debe a Heródoto (1.400 AC) quien menciona que en Egipto trabajaban unos técnicos llamados "estiradores de cuerdas", que utilizaban cuerdas de longitudes conocidas con las que se encargaban de replantear los límites de las propiedades después de las crecidas del río Nilo, asignando a cada agricultor el área que le correspondía tal cual había sido relevada previo a la crecida. Esto permitía mantener el funcionamiento del motor principal de la economía como era la agricultura.

Las anotaciones respecto de las áreas que les correspondían a los agricultores no sólo servían para el replanteo de sus propiedades sino para el cobro de un gravamen proporcional a la producción potencial de la parcela (Larsson, 1991). Este ejercicio plantea los dos fines principales de los catastros, el primero registrar la localización de las propiedades y sus respectivos usos, el segundo, cobrar impuestos a la propiedad basado en tamaño de la propiedad y el uso de la misma. Estos principios predominan en la actualidad.

El valor del suelo es un concepto principalmente económico. Sin embargo, al enumerar los atributos del suelo existe una serie de puntos de coincidencia entre la economía y la geografía (Appraisal Institute, 2013). Estos son:

- Cada predio de suelo es único en su localización y composición
- El suelo es físicamente inamovible
- El suelo es un bien durable
- El suelo es muy útil para la sociedad

Uno de los factores que influye en el valor del suelo es el uso al que se dedica éste. En países donde la tenencia y renta del suelo no es libre, es el gobierno quien dictamina el uso de éste. En economías de libre mercado, el uso del suelo es regularizado por una serie de regulaciones.

En esta discusión es imperativo comprender el principio de “valor de mercado”. El valor de mercado para un bien representa el precio monetario por el cual se decide intercambiar el bien entre el vendedor y el comprador (Appraisal Institute, 2013). Este intercambio se debe efectuar en un ambiente de pleno conocimiento de las características y condiciones principales del bien que se desea intercambiar. Para que la venta del bien inmueble establezca el valor de mercado del mismo, las partes envueltas deben reconocer el valor actual del inmueble al igual que el uso potencial del mismo.

1.2 Evolución de la contribución sobre la propiedad inmueble en Puerto Rico.

Es prioritaria una revisión profunda al régimen de contribución sobre la propiedad inmueble en Puerto Rico. La última tasación general de propiedades inmuebles se realizó hace poco más de medio siglo. Los criterios y procesos adoptados para los procesos de valoración de inmuebles datan de 1957, por lo tanto, al tasar nuevos inmuebles, el Centro de Recaudación de Ingresos Municipales (CRIM)¹ utiliza métodos de aproximación basados en premisas que hace décadas dejaron de ser válidas.

¹ El Centro de Recaudación de Ingresos Municipales (CRIM por sus siglas en español) es una entidad municipal creada en el 1991 como parte del proceso de Reforma Municipal para acercar a los municipios a la administración de sus principales fuentes de ingresos, vitales para el financiamiento de los servicios que ofrecen y de sus operaciones. Asimismo, con la creación del CRIM se delegó a este, en representación de los

Hasta antes del 1991, con la aprobación de las leyes de reforma municipal, las funciones y poderes relativos a la contribución sobre la propiedad eran controlados y administrados por el Secretario de Hacienda siguiendo las directrices del “Código Político de Puerto Rico de 1902”. En el artículo 302 de dicho código se establece que el tasador deberá decidir el valor del inmueble "en su valor real y efectivo, sin tener en cuenta una venta forzada". Para la valoración de la propiedad se entregaba una planilla al dueño del inmueble donde enumeraba y describía las características del mismo. Basado en lo declarado en la planilla el tasador procedía a valorar el inmueble. De no contar con una planilla con datos sobre el inmueble, el tasador procedía a valorar el mismo según “los mejores informes que pueda tener” (Código Político de Puerto Rico de 1902). Esto consistía de obtener información del inmueble por parte de los vecinos del mismo. Este proceso no estaba sujeto a ninguna metodología ni procedimiento formal. Surgían así procesos de valoración sin objetividad donde el resultado en ocasiones era que inmuebles de características similares presentaban valores diferentes.

El 9 de mayo de 1947 se aprueba la Ley núm. 117 conocida como “Ley de Catastro, Clasificación y Tasación de la Propiedad”. El propósito de esta ley fue facultar al Tesorero de Puerto Rico a que adoptara un método de tasación que aplicara a los distintos tipos de propiedad existentes en Puerto Rico de manera que la misma se diera de manera uniforme. De esa forma se buscaba corregir el problema de falta de objetividad que presentaba el modelo establecido en el Código de Puerto Rico de 1902.

El Código de Puerto Rico de 1902 establecía la necesidad de tasar los bienes inmuebles, sin embargo, no establecía la necesidad de crear un catastro para registrar esos bienes. La Ley 117 le requirió al Tesorero de Puerto Rico realizar el catastro o inventario de toda la propiedad inmueble de Puerto Rico, además de clasificar y tasar esta. Instituyó la ley que en el primer paso se debía dar la clasificación de las propiedades por clase y tipo, con el objetivo de tasar las propiedades basado en la categoría a la que perteneciera y luego establecer normas de valoración para estas. Con el fin de lograr tasaciones uniformes para los distintos tipos de propiedad el método seleccionado debía contener normas de valoración con "exactitud y detalles científicos" (Ley Número 117 de 1947). Entiéndase por "científico" que las normas debían contener las "exigencias de precisión y objetividad propias de la metodología de las ciencias" (Ley Número 117 de 1947). Además, los métodos seleccionados

municipios la administración de la Ley de Contribución Municipal sobre la Propiedad de 1991. El CRIM rinde servicios fiscales a favor de los municipios y como organismo municipal tiene la responsabilidad de notificar, tasar, recaudar y distribuir los fondos públicos provenientes de distintas fuentes dispuestas por ley. Además, es responsable de poner y mantener al día el Catastro de Propiedad Inmueble de Puerto Rico.

deberían ser métodos "reconocidos en materia de valoración". Se entiende así que se adoptó un método de tasación para cada clase o tipo de propiedad.

Para la valoración de las propiedades inmuebles, el Tesorero desarrolló valores unitarios para distintos componentes del inmueble. El modelo de valoración desarrollado consideraba factores como el acceso del inmueble, su localización, uso y servicio de electricidad, agua y sistema de alcantarillados. En cuanto a la tipología edificatoria, las estructuras fueron clasificadas por clases y subclases utilizando como criterio para esta clasificación los materiales que se utilizaron en la construcción junto a la calidad de estos y el uso al que sería destinada la estructura. Esta metodología de tasación, de asignar unos valores unitarios, es lo que se denominó como “Proyecto de Tasación Científica”. Los valores que se establecieron se basaron en el costo de reemplazo del 1951. En otras palabras, asignar un valor basado en la cantidad de dinero necesaria para re-construir una propiedad inmueble igual a la existente utilizando el valor de los materiales de construcción del año 1951.

Con la adopción de la Constitución del Estado Libre Asociado de Puerto Rico el 25 de julio de 1952, se establece el Departamento de Hacienda como un Departamento Ejecutivo. Con este cambio el Tesorero de Puerto Rico se convierte en el Secretario de Hacienda de Puerto Rico. En el año fiscal 1958 – 1959 se realiza una revaloración de la propiedad mediante un reajuste en los valores unitarios establecidos en el 1951, utilizando los criterios, normas, principios e instrumentos implantados en el 1951 – 1952. Se realizaron estudios para determinar el valor en el mercado de las propiedades al 1 de enero de 1957. Es entonces que comienza el proyecto de revaloración de los bienes inmuebles acompañado de una reestructuración en la tasa contributiva.

Los suelos urbanos fueron clasificados aplicando factores que afectaban el valor conforme a la variación en los tipos de uso. En el caso de los suelos rurales se crearon valores mediante la utilización de tres métodos de análisis de tasación: análisis de venta, análisis de renta y análisis de ingresos generados por fincas agrícolas. Las edificaciones o estructuras en estos se clasificaron de acuerdo al uso para el cual fueron diseñados y el tipo de construcción determinado por los materiales utilizados y la calidad de los mismos.

Este sistema de valoración se mantuvo intacto hasta el año 1991 cuando como parte del proceso de reforma municipal de Puerto Rico se aprueba la Ley número 83 de 30 de agosto de 1991 mejor conocida como “Ley de Contribución Municipal sobre la Propiedad de 1991”. Esta ley impone la creación del “Centro de Recaudación de Ingresos Municipales” (CRIM).

Al CRIM se le delega la administración de las contribuciones municipales sobre la propiedad que antes ejercía el Secretario del Departamento de Hacienda. A partir de ese momento el CRIM es la entidad responsable de clasificar, tasar, notificar y cobrar las contribuciones mueble e inmueble en todo Puerto Rico.

La Ley 83 dispone que "El Centro de Recaudación de Ingresos Municipales procederá a hacer un plan que permita la revisión constante de la propiedad inmueble, a fin de mantenerla al día". Este aspecto tiene su origen en el artículo 9 de la Ley 117 de 1947, la cual ordenaba al Tesorero "hacer un plan que permita la revisión constante de la tasación a fin de mantener la valoración de la propiedad lo más en armonía posible con el valor de las propiedades según éste vaya variando". En otras palabras, el propósito de estas disposiciones era que los valores adoptados en el 1957 para los inmuebles, se fueran actualizando con los cambios del mercado de manera que se mantuviera el valor al día. Esto no ocurrió durante el tiempo en que Hacienda administró dicha ley y al momento de esta redacción no ha ocurrido con el CRIM. Esto hace que la agencia aplique impuestos basado en valores y condiciones que corresponden al año 1957.

De manera contradictoria, la misma Ley 83 dispone que el CRIM no puede realizar cambios en los métodos de tasación y hacer una nueva tasación hasta tanto haya tasado sustancialmente todas las propiedades en Puerto Rico. Esto aún no se ha alcanzado. La falta de coordinación existente entre las agencias gubernamentales con intereses en el proceso de otorgación de permisos, registro de la propiedad y catastro hace que el CRIM no tenga al día el número de inmuebles hábiles para el proceso de valoración. En el caso del Registro de la Propiedad de Puerto Rico y el CRIM (el Catastro) la inscripción de propiedades por parte del ciudadano es voluntaria. Durante la década del 1970 se creó un esfuerzo entre el Registro de la Propiedad de Puerto Rico y el Secretario del Departamento de Hacienda para identificar inmuebles que estuvieran registrados en el primero y no en el segundo. Fueron muchos los inmuebles para los que se identificaron planos. Sin embargo, la falta de personal en el Departamento de Hacienda hizo que el trabajo fluyera lentamente. En muchos casos los planos desaparecieron, haciendo difícil el proceso de actualización. El peor de los casos es de propiedades creadas por procesos de lotificación (en especial las zonas rurales) que nunca se inscribieron ni en el Registro de la Propiedad de Puerto Rico ni en el Departamento de Hacienda. Estos simplemente son inmuebles que existen y llevan décadas sin pagar impuestos. Según estimados del CRIM esta cifra debe rondar 140.000 viviendas sujetas a contribución.

Para mejorar esta situación algunos municipios han sido diligentes y han realizado el trabajo que le corresponde al CRIM al identificar y tasar estos inmuebles utilizando diversas fuentes de información. Esto se ha logrado por medio de convenios entre los municipios y el CRIM donde el último acepta el trabajo realizado por los municipios para actualizar sus bases de datos. Además, han identificado nuevas construcciones dentro de sus límites territoriales y han asignado sus propios tasadores para valorar las propiedades. Los municipios con menos recursos dependen exclusivamente del CRIM.

Bajo la ley actual, el dueño de una propiedad residencial sólo paga contribución sobre su residencia principal en tanto y en cuanto la tasación de la misma exceda 15.000 dólares (al valor del año 1957), que es el monto de la exoneración estatutaria. En el caso de propiedad residencial localizada en cualquier municipio que imponía contribuciones sobre la propiedad al 30 de agosto de 1991 (fecha en que aprueba Ley de Reforma Municipal y la que crea el CRIM), y con respecto a la cual se hubiese solicitado exoneración en o antes del 1 de enero de 1992, el gobierno central está obligado a devolver a los municipios la contribución sobre la propiedad atribuible a dicha exoneración estatutaria, hasta un máximo de 2%. A los valores actuales de las residencias, la exoneración podría representar en muchos casos un valor en el mercado de entre 200.000 y 250.000 dólares, dependiendo de la localización de la residencia.

Según informe de auditoría de la Oficina del Contralor del Puerto Rico, al 29 de marzo de 2012, en el archivo de contribuyentes de propiedad inmueble de la base de datos de la agencia existían 1.365.483 registros de contribuyentes de propiedad inmueble. El valor aproximado tasado de estas alcanza la cifra de 170.000 millones dólares. Con solo revaluar y actualizar los valores de los inmuebles se podría generar aproximadamente 550 millones de dólares adicionales por año. Sin embargo, dado que el proceso de revaloración es uno de poco agrado a los ciudadanos y que puede implicar un costo político para los gobernantes que lo impongan, se han concentrado más en la identificación de los inmuebles no tasados antes que realizar un proceso de revaloración.

Ante la crisis económica, en el año 2008, el gobierno de Puerto Rico designó un comité para analizar la situación del país y proponer alternativas de intervención. El grupo llamado “Consejo Asesor de Reconstrucción Económica y Fiscal” rindió un informe donde, entre varios aspectos, recomendaron que comenzando con el año fiscal 2009 - 2010, se impusiera una sobretasa de contribución estatal igual al 100% del monto de la contribución sobre la propiedad inmueble que cada dueño tributa al CRIM. El gobierno aceptó la recomendación y

por medio de la Ley 7 del 9 de marzo del 2009, mejor conocida como “Ley Especial Declarando Estado de Emergencia Fiscal y Estableciendo Plan Integral de Estabilización Fiscal para Salvar el Crédito de Puerto Rico” se impuso esta medida. El problema es que la misma es una revalorización arbitraria de las propiedades donde la sobretasa resulta ser diez veces el valor tributable. Por ejemplo, un ciudadano que cuenta con una exoneración contributiva de 15.000 dólares y el valor tributable de su inmueble era de 24.090 dólares con la contribución especial la exoneración subió a 150.000 dólares y el valor tributable a 240.900 dólares. Esto no representa un proceso científico ni responsable de establecer el valor de un inmueble.

1.3 El modelo de Puerto Rico comparado

Para comprender la condición actual del modelo de valoración inmueble con fines contributivos en Puerto Rico es importante compararlo con otros modelos desarrollados en el mundo. Como se ha discutido al momento, en Puerto Rico el valor inmueble se fundamenta en los valores adoptados en el año 1957. Estos valores se dividen en el valor del suelo y el valor de la estructura. Los ciudadanos cuentan con una exoneración contributiva de 15.000 dólares del valor tributable del inmueble que utilice como vivienda principal. Una vez se suman los valores del suelo y la estructura y se resta la exoneración contributiva se obtiene el valor sujeto a tributación. Cada municipio tiene la autonomía de establecer el tipo contributivo que aplica a los inmuebles de su territorio el año fiscal corriente. Para el año fiscal 2016 - 2017 los tipos contributivos fluctuaban entre 8,03% y 1,83%.

En el caso de los **Estados Unidos de América** la entidad encargada de realizar la valoración del inmueble y el cobro de impuestos por su valor es el condado (“county” en inglés)². Los estados adoptan las leyes que regulan los procesos de valoración de inmuebles y cobro de impuestos para sus territorios y fundamentados en estas leyes los condados adoptan sus propios modelos de valoración (Harris y Moore, 2013). Con el objetivo contributivo, los condados gravan los bienes inmuebles principalmente imponiendo una tasa al valor justo de mercado de la propiedad (Commeaux, 2015). El valor del inmueble se establece principalmente utilizando el valor del mercado del mismo. Como método común se utiliza la

² Un condado en los Estados Unidos es un nivel local de gobierno creado como una subdivisión de un estado. Este puede ser creado por el gobierno estatal, el federal o por el gobierno territorial como una subdivisión de un territorio. La unidad comparable en Puerto Rico es el municipio

valoración en masa de inmuebles utilizando métodos de regresión y valores de ventas inmuebles realizadas en el año previo al ejercicio de valoración. El período de actualización de los valores de inmuebles puede variar de estado a estado. En algunos casos se deben actualizar anualmente, en otros casos se da un espacio de tiempo más amplio (en Texas, tres años: en Luisiana, cuatro años (Commeaux, 2015).

Algunos estados tienen requisitos de uniformidad, asegurando una igualdad en el cobro en todo el estado (Commeaux, 2015). Con estos establecen límites que controlan el grado en que los impuestos sobre la propiedad pueden aumentar en un año determinado. Algunos estados con altos impuestos a la propiedad, como New Hampshire y Texas, dependen en gran medida de estos impuestos en lugar de otras medidas impositivas (Harris y Moore, 2013). Al año 2016, Nueva Jersey contaba con la tasa más alta efectiva en 2,38% seguido por Illinois (2,32%), New Hampshire (2,15%) y Connecticut (1,98%). En el otro extremo del espectro, Hawái tiene la tasa efectiva más baja en 0,28%, seguido por Alabama (0,43%), Luisiana (0,51%), y Delaware (0,55%).

En el caso de **Argentina** el impuesto inmobiliario es establecido por las provincias. Estas están encargadas tanto de legislar como de recaudar los impuestos por este concepto (Molinatti, 2016). Las provincias han delegado en los municipios la tarea tributaria. Aunque esto no ha ocurrido de manera uniforme. En algunas provincias se han otorgado facultades impositivas expresas sobre la propiedad inmueble, en otros casos solo se les permite realizar el recaudo (Molinatti, 2016). En cuanto a las tasas, no existe un sistema de tarifas uniforme. Se pueden identificar tasas progresivas (mayor valor, mayor tasa), en otros casos las tasas son por escalas de valor (Benegas, 2001). En algunos casos los lotes baldíos cuentan con unas tasas mayores en unas ciudades que en otras. En todos los casos las cuotas son aprobadas por el poder legislativo de las provincias y en las ordenanzas tarifarias municipales. El método de valoración lo establece la jurisdicción que administra el tributo. Se parte del principio que el método adoptado logre un precio justo (Molinatti, 2016). En términos de la valoración, la misma se obtiene sumando los valores del suelo y de las estructuras. Para establecer el valor del suelo se obtienen valores de suelo libre de mejoras tomando muestras de ventas en el mercado de predios con estas condiciones (Molinatti, 2016). Para el valor de la construcción se utiliza el método de costo de reposición. Este se calcula tomando en consideración la antigüedad y estado de conservación de la construcción (Molinatti, 2016).

En el caso de **Brasil** el impuesto que incide sobre los inmuebles urbanos se conoce como el IPTU (Imposto sobre a Propriedade Predial e Territorial Urbana). Este representa la segunda fuente tributaria de los gobiernos municipales (De Cesare, 2016). A nivel nacional se cuenta con el Código Tributario Nacional (CTN). En éste se definen las normas tributarias y las disposiciones básicas de los diferentes tributos en el país. La base del cálculo que utiliza el IPTU los establece el CTN (de Cesare, 2016). El IPTU solo maneja el tributo de inmuebles urbanos, los inmuebles rurales son de competencia del gobierno federal. Las tasas se establecen por medio de leyes municipales. Las cuotas pueden variar de municipio a municipio entre un 0,50% y un 7% (de Cesare, 2016). En términos de la valoración de los inmuebles, no existe una norma específica que tengan que seguir los municipios. La única norma general la establece la oficina de Catastro Territorial Multifinalitario al indicar que los trabajos de valoración de inmuebles con fines tributarios se realicen de acuerdo a las normas de la Asociación Brasileña de Normas Técnicas (Associação Brasileira de Normas Técnicas), así como que el ciclo de valoración sea como máximo cada cuatro años (de Cesare, 2016). Estas instrucciones además discuten factores de calidad de los valores de catastro y estimados de valor como es el Coeficiente de Dispersión.

De los impuestos cobrados en **Chile**, el del impuesto predial es de los más antiguos teniendo su origen en 1916. La regulación del impuesto predial en este país está a cargo del Servicio de Impuestos Interno (SII), entidad dependiente del Ministerio de Hacienda. Esta entidad es responsable de determinar la base del impuesto y mantener actualizado el catastro (Orrego, 2016). El proceso de cobro del impuesto recae en la Tesorería General de la República, entidad también parte del Ministerio de Hacienda. En el caso chileno los municipios solo obtienen el beneficio económico que genera este impuesto. No tienen participación en ninguna de las fases relacionadas al cobro de este impuesto (Orrego, 2007).

En torno a las tasas para inmuebles se establecen tres categorías; Bienes raíces agrícolas, Bienes raíces no agrícolas, Bienes raíces no agrícolas destinados a la habitación (Orrego, 2007). Por esta última, que incluye los bienes inmuebles, se establece un 1,2 por ciento al año, en la parte de la base imponible que no exceda de \$37.526.739 (Pesos Chilenos) al 1 de enero de 2003; y 1,4 por ciento al año, en la parte de la base imponible que exceda del monto señalado (Orrego, 2016).

En el año 2014 entra en vigencia un proceso de actualización de valores inmuebles que busca establecer la estructura de datos inmuebles basado en información de transacciones realizadas

en tiempos recientes. La información se obtiene electrónicamente desde las Notarías y Conservadores de Bienes Raíces y de tasaciones comerciales de bancos e instituciones financieras, remitidas por la vía electrónica (Orrego, 2016). Cada comuna del país se divide en Áreas Homogéneas (AH), que corresponden a sectores que presentan características urbanas comunes. Para estas se toman muestras de ventas inmobiliarias que por medio de modelos estadísticos permite establecer los valores unitarios del inmueble para el Área Homogénea. Para cada AH existe una ficha descriptiva que contiene información de las características consideradas en la determinación de su valor de terreno (Orrego, 2016).

En el caso de **Colombia** el ejercicio de valoración y cobro de inmuebles es uno dividido. Los avalúos de inmuebles catastrales son responsabilidad del Instituto Geográfico Agustín Codazzi (IGAC) y la recaudación y cobro del impuesto es de competencia municipal (Montaña, 2016). Además, los municipios tienen la facultad de imponer medidas cautelares como el embargo y remate de inmuebles (Iregui, et al, 2004). La función de establecer las tasas impositivas recae en los municipios. Estos pueden establecer tasas que varían desde el 0,5 por ciento al 1,6 por ciento. Entre los criterios utilizados para establecer la tarifa no solo está el valor catastral, también se considera el estrato social en que se encuentra el inmueble, los usos del suelo del sector urbano y la antigüedad de la formación o actualización del catastro (Montaña, 2016). Para los predios no urbanizados o sin ningún tipo de construcción el municipio puede establecer una tarifa mayor al 1,6 por ciento siempre y cuando nunca exceda el 3,3 por ciento. En la actualidad las ciudades con las tasas más altas son Medellín y Cali. La autonomía en el establecimiento de las tasas se puede evidenciar con el caso de Bogotá. En esta ciudad las residencias que tengan un valor catastral sobre los 229 millones de pesos colombianos (\$76.000 / €72.000) la tasa es del 0,95 por ciento del valor. Si el valor catastral está por debajo de esta cifra se impone una tasa del 0,7%.

El IGAC debe establecer los valores inmuebles siguiendo la dinámica del mercado inmobiliario. Por tal razón debe realizar las revisiones de estos valores en un período máximo de cinco años. En el proceso de actualización de los valores inmuebles, el resultado del valor no puede ser inferior al 60 por ciento de su valor de mercado (Montaña, 2016). Parecido al caso de Chile, en Colombia el avalúo de los inmuebles se realiza en zonas homogéneas geoeconómicas.

España cuenta con un modelo similar al colombiano donde la gestión del impuesto de inmuebles es compartida entre el Estado y los Ayuntamientos. En su caso este impuesto es

conocido como el IBI o Impuesto sobre Bienes Inmuebles. El hecho imponible está constituido por la propiedad de los bienes inmuebles de naturaleza rústica y urbana localizados en el término municipal, o por la titularidad de un derecho real de usufructo o de superficie, o de la de una concesión administrativa sobre dichos bienes o sobre los servicios públicos a los que estén afectados, y grava el valor de los referidos inmuebles (Miranda, 2012).

El valor del inmueble sobre el cual se establecerá la tasa se determina a partir de los datos que contiene el catastro inmobiliario. Este se compone del valor catastral del suelo y el valor de las edificaciones (Palomo y Müller, 2010). Como regla general este valor no puede superar al del mercado. Estos valores pueden ser actualizados aplicando diversos coeficientes para cada clase de inmueble. Las tasas aplicables dependen del tipo de inmueble. Los inmuebles urbanos se aplica una tasa de entre el 0,4% y el 1,10%, para los inmuebles rústicos una tasa de entre el 0,3% y el 0,90% y para los inmuebles de características especiales una tasa de 0,6% (Miranda, 2012).

La Dirección General del Catastro indica que se puede determinar un nuevo valor catastral para la totalidad de los bienes inmuebles de un municipio por medio de procedimientos de valoración colectiva de carácter general (PVCCG). Estos se inician previa manifestación del interés por el Ayuntamiento correspondiente, al detectarse diferencias sustanciales entre los valores de mercado actuales y los que sirvieron de base para la determinación de los valores catastrales vigentes. Los procedimientos de valoración colectiva de carácter general requieren la aprobación de una ponencia de valores total. Las Ponencias de Valores totales recogen los criterios, módulos de valoración y demás elementos precisos para llevar a cabo la determinación del valor catastral de cada bien inmueble. Son aprobadas por el Director General del Catastro, previo informe del Ayuntamiento correspondiente, publicándose el acuerdo de aprobación en la Sede Electrónica del Catastro. Los valores catastrales resultantes de estos procedimientos se notifican individualmente a los titulares catastrales y tendrán efectividad el día 1 de enero del año siguiente (Dirección General del Catastro, 2016).

La actualización de los valores catastrales también se puede realizar por medio de coeficientes (Palomo y Müller, 2010). Con esto se busca aproximar los valores catastrales al valor actual de mercado de los bienes inmuebles. Estos coeficientes pueden ser diferentes para cada grupo de municipios o para cada clase de inmueble. En todo caso el municipio debe

manifestar diferencias sustanciales entre los valores de mercado y los que sirvieron de base para la determinación de los valores catastrales vigentes.

Es importante indicar que los seis modelos analizados cuentan con procesos de impugnación por parte de los ciudadanos que no estén de acuerdo con el valor de inmuebles establecido para su propiedad. Además, cuentan todos con listado de excepciones de costo que varían desde excepción a facilidades gubernamentales, residencias principales y ciudadanos que superen ciertos límites de edad.

1.4 El Problema del Modelo Actual

De la situación planteada se puede concluir que los criterios que se utilizan para la valoración y cobro de contribuciones sobre los inmuebles en Puerto Rico no permiten que el sistema sea justo ni equitativo ya que estos datan de un procedimiento establecido en 1951 y ajustado en 1957, esto es, sobre cincuenta años sin revisión de los mismos. Además, dado que no se realiza un proceso de actualización de los datos y los dueños de inmuebles no informan de cambios estructurales en términos de mejoras como ampliaciones, construcción de terrazas o de piscinas, estos cambios estructurales del inmueble no están contemplados en el valor de tasación asignado. Se tributa basado en el valor asignado originalmente al inmueble una vez se inscribió y se tasó por primera vez con las tasas establecidas en 1957.

Los estudios de valoración contemporáneos establecen, entre otros, que la estimación del valor del suelo está condicionada por los aspectos espaciales de los factores que permiten su determinación. La misma ley 117, en su época, reconocía que en la fórmula para estimar el valor se deberían considerar factores como acceso, localización, uso y servicio de electricidad, agua y sistema alcantarillado. Algunos de estos factores son de naturaleza espacial, un cálculo que en los tiempos actuales debe ser fácil de establecer dadas las herramientas informáticas con que se cuenta, en especial los Sistemas de Información Geográfica. Muchos estudios reconocen que el valor de una residencia, más que un efecto de aspectos espaciales, es un efecto directo de condiciones espacio - temporales (Quercia et al., 2000).

Por lo tanto, el problema que se observa es la ausencia de un proceso de actualización de los valores de inmuebles en Puerto Rico por parte de la agencia encargada. A pesar que las leyes

creadas para establecer los procesos de valoración de inmuebles dan espacio para que esta actualización ocurriera, la realidad es que nunca se llevó a cabo la misma. La cantidad de unidades de vivienda en Puerto Rico en el año 1950 totalizaban 453.572. En la actualidad esa cantidad totaliza 1.636.946. El proceso de valoración adoptado en el 1950 implicaba la visita de cada inmueble. Tomando la diferencia en inmuebles, este ejercicio resulta ser más complejo en la actualidad. Además, se debe tomar en consideración la significativa inversión económica que implica esta tarea siguiendo este método. En la actualidad se cuenta con herramientas y tecnologías que en la década del 1950 no existían, los cuales pueden ayudar en los procesos de adquisición de información. Dentro de estos se encuentra los Sistemas de Información Geográfica (SIG).

Se pueden citar varias investigaciones donde se utilizan los SIG para estudiar algún aspecto de valores residenciales (Bitter et al, 2007; Yaolin, 2006; Anselin y Lozano García, 2008; Lu et al., 2014; Larson 2015;) Uno de los precursores en el uso de SIG para valoración de inmuebles es Ayse Can. Éste desarrolla un marco analítico para el uso de Sistemas de Información Geográfica en estudios del mercado hipotecario de residencias donde utiliza características recopiladas de los inmuebles para determinar el valor de los mismos. Para lograr esto, Can utiliza la estimación hedónica como fundamento para su hipótesis³. Integra en el SIG datos sobre cuatro grupos de factores con los cuales establece los valores de residencias por vecindarios (Can 1998).

La metodología de Can es ejemplo de la capacidad de determinar el valor de los inmuebles con datos sobre características de los mismos. Sin embargo, no considera el hecho que estas características pueden cambiar o que su efecto en el valor del inmueble por medio de la técnica hedónica puede variar en el tiempo. Este factor de cambio en el tiempo de las variables está ausente, tanto de las investigaciones científicas, como de las metodologías adoptadas por diversos países con el fin último de estimar la valoración de inmuebles. Se convierte ésta en una debilidad de estos modelos.

Al comparar los modelos adoptados por diversos países, se pudo observar que estos acogen diferentes períodos de actualización de los valores de inmuebles. En ninguno de los casos analizados se presenta una justificación del período de tiempo adoptado para esta tarea. Esta situación genera dudas como pueden ser; ¿puede el período de tiempo para actualizar el valor

³ La técnica de valoración por el método hedónico se fundamenta en la idea que el valor de una propiedad es función no solo del valor de las características estructurales que comprenden el inmueble sino también del valor de externalidades del inmueble.

del inmueble ser mayor al adoptado? ¿es justo este valor cuando se compara con aquellos del mercado? o ¿qué sucede cuando se cuenta con un valor adoptado y surgen condiciones en el mercado que hacen que el valor de mercado de los inmuebles se reduzca? El ejercicio de realizar la actualización de los valores inmuebles puede llegar a ser altamente costoso. Comprender el período de validez de las valoraciones podría permitir adoptar intervalos de tiempo para la actualización más ajustados a la realidad de cada país. Por tal razón, adoptar un proceso para entender este aspecto de la valoración se convierte en una de las metas principales de la presente investigación.

La investigación que aquí se presenta busca seleccionar un método que permita realizar esta actualización de los valores inmuebles de manera ágil y precisa comprendiendo a la vez el cambio temporal de diversas valoraciones inmuebles con el objetivo de adoptar un período razonable de actualización de los valores inmuebles.

1.5 Como intervenir en la Situación Actual: Objetivos de la Investigación

La idea principal de la investigación es construir un sistema para establecer valores de inmuebles residenciales que permita el cobro de impuestos por propiedad inmueble acorde al mercado inmobiliario y se establezca el período de tiempo por el cual esa valoración es vigente. Esto permitirá sustituir el modelo actual y actualizar los criterios de valoración que datan de la década del 1950. En el sistema propuesto se interrelacionan diversos factores de índole espacial y temporal que inciden en la valoración del suelo.

De la idea principal de la investigación se derivan una serie de objetivos:

- a. Adoptar un método que permita establecer una base de valores de inmuebles residenciales con fines contributivos más precisa y de ágil actualización.
- b. Establecer la validez temporal de los valores de inmuebles residenciales como base para el componente contributivo del inmueble.
- c. Utilizar Sistemas de Información Geográfica para modelar los valores de inmuebles residenciales y establecer la validez temporal de los mismos.

d. Generar sugerencias que ayuden a las agencias relacionadas implantar una base de valores de inmuebles residenciales actualizado que permita el cobro de impuestos por propiedad inmueble de manera justa y real.

Establecer una base de valores de inmuebles residenciales con fines contributivos más precisa (primer objetivo) se fundamentó en el concepto de la valoración hedónica. La utilización de la valoración hedónica se apoya en el uso de técnicas de regresión, siendo la más común los Mínimos Cuadrados Ordinarios. En años recientes se ha utilizado como alternativa a los Mínimos Cuadrados Ordinarios la técnica de Regresión Geográficamente Ponderada. Ésta es una técnica de regresión espacial que proporciona un modelo local de la variable o proceso que intenta entender al ajustar una ecuación de regresión a cada entidad en el grupo de datos utilizando. Los Sistemas de Información Geográfica facilitan tanto el almacenamiento de grandes volúmenes de información georreferenciada como su tratamiento estadístico o geoestadístico. Así, utilizando éstos, se realizó una comparación entre ambas técnicas. Se desarrolló un proceso fundamentado en un grupo de pruebas estadísticas que permitió seleccionar la más adecuada de las dos.

Para cumplir con el establecimiento de un método que permitiera determinar la validez temporal de los valores de inmuebles residenciales como base para el componente contributivo del inmueble (segundo objetivo) se seleccionaron datos de ventas de inmuebles para la zona de estudio que van del 2005 al 2009. La idea de seleccionar datos de este intervalo de tiempo es que el modelo se enfrente al comportamiento inmediato del mercado previo y pasado al estallido de la burbuja inmobiliaria por el que atravesó Puerto Rico en el año 2007 (Coy, 2006; Krugman, 2009; Shiller, 2011). Ejecutar el modelo con datos de un solo año tan solo reflejaría las condiciones del mercado del año seleccionado, no reflejaría los cambios en el tiempo del mismo. Por tal razón la selección de estos años.

Para corroborar la efectividad en la predicción de los valores inmuebles por parte de los modelos seleccionados, se compararon estos con valores residenciales de ventas realizadas entre los años 2010 al 2014. Esto permitió medir la efectividad del modelo buscando observar el comportamiento de los datos en el período de efectividad de las predicciones, esto con relación a la ejecución del mercado inmobiliario.

El estudio de la vigencia de los valores de predicción se fundamentó en tres pruebas; el Error Porcentual Absoluto Medio y el Error Absoluto Medio y el Coeficiente de Dispersión (COD). El Error Porcentual Absoluto Medio es un indicador del desempeño del pronóstico de

demanda que mide el tamaño del error (absoluto) en términos porcentuales. El hecho que se estime una magnitud del error porcentual lo hace un indicador frecuentemente utilizado por los encargados de elaborar pronósticos debido a su fácil interpretación. El Error Porcentual Absoluto Medio da una idea clara del desempeño de los datos analizados, pero su significado en términos monetarios da una idea más clara del impacto en dólares que presentan estas relaciones. El Coeficiente de Dispersión es una medida de la uniformidad de la valoración de todas las propiedades de una ciudad. Mide la desviación media entre los precios de venta de las propiedades vendidas recientemente de la media de valoración de inmuebles de la ciudad. La Asociación Internacional de Oficiales Evaluadores (International Association of Assessing Officers o IAAO por sus siglas en inglés) publica normas de uniformidad utilizando el coeficiente de dispersión. Se utilizaron estas normas para analizar los resultados obtenidos. El Error Porcentual Absoluto Medio es un indicador del desempeño del pronóstico de demanda que mide el tamaño del error (absoluto) en términos porcentuales. El hecho que se estime una magnitud del error porcentual lo hace un indicador frecuentemente utilizado por los encargados de elaborar pronósticos debido a su fácil interpretación. El Error Porcentual Absoluto Medio da una idea clara del desempeño de los datos analizados, pero su significado en términos monetarios da una idea más clara del impacto económico que presentan estas relaciones.

La modelación y análisis de los datos se realizó con el programa ArcGIS. Este programado permitió no solo modelar el aspecto espacial de los datos de ventas inmuebles, además permitió realizar las pruebas utilizando los Mínimos Cuadrados Ordinarios y la Regresión Geográficamente Ponderada (tercer objetivo).

Basado en la experiencia de la investigación, los resultados obtenidos y la experiencia de modelos adoptados en otros países, se generaron sugerencias que pueden ser adoptadas localmente y que permiten adoptar un modelo actualizado que viabilice el cobro de impuestos por propiedad inmueble de manera justa y real (cuarto objetivo).

1.6 Estructura de la Investigación

La tesis se presenta en ocho capítulos. Luego de la introducción a la investigación en el primer capítulo se procede a desarrollar en el segundo capítulo una discusión de los conceptos teóricos referentes al valor del suelo y las teorías y métodos existentes para la

valoración de inmuebles residenciales. Luego pasa a discutir los modelos espacio-temporales que han predominado la investigación sobre el tema en los pasados años. En el tercer capítulo se comienza discutiendo el concepto de la valoración hedónica basado en los modelos de regresión y luego se pasa a discutir los fundamentos de la Regresión Geográficamente Ponderada. Ambas son discutidas bajo la idea de utilizar estas para el análisis de valores residenciales inmuebles.

En el cuarto capítulo se describe la zona de estudio y en el quinto se discute todo el análisis exploratorio de datos que fue necesario realizar para seleccionar el modelo de datos idóneo. Con el modelo de datos seleccionado se procedió a aplicar las técnicas de regresión con la técnica de los Mínimos Cuadrados Ordinarios y Regresión Geográficamente Ponderada. Los resultados, la comparación de estos y la selección de la técnica que mejor modela los datos se discutió en el capítulo sexto. Los resultados del modelo fueron validados utilizando diversas pruebas estadísticas predominando el Coeficiente de Dispersión. Esto se discute en el séptimo capítulo. Finalmente, con los resultados obtenidos se procedió a realizar un planteamiento final y concluir la mejor forma de manejar el problema planteado desde el inicio de la investigación. Se tomó en consideración los modelos de valoración de inmuebles con fines contributivos discutidos en el primer capítulo y a la luz de sus experiencias y lo observado en los resultados de la presente investigación se realizaron sugerencias para mejorar la situación en Puerto Rico.

Capítulo 2

Valoración Inmobiliaria: Conceptos Básicos

2. Valoración Inmobiliaria: Conceptos Básicos

2.1 El suelo como mercancía

De la discusión previa se puede concluir que el valor de los inmuebles residenciales se compone de dos factores, primero el factor suelo y segundo las mejoras que se realicen en el mismo. El valor del suelo se refiere a los ingresos devengados del suelo en el proceso de producción (Kauko, 2014). Se puede entender que la producción está condicionada a las mejoras que se den en el suelo. En el caso que el suelo no se utilice productivamente, el valor del mismo se basará en el costo de oportunidad de no colocar el mismo en el proceso de producción (Kauko, 2014; Appraisal Institute, 2013). Se entiende por costo de oportunidad la pérdida de oportunidades que surgen al momento de asignar al suelo lo que se entiende es la alternativa de mayor y mejor uso al que podría ser sometido el mismo (Appraisal Institute, 2013). A este planteamiento, se añade el concepto temporal, al establecer que estas mejoras y el ingreso que representan, pueden expresarse como ingresos por concepto de alquiler / valor que ocurre en términos variables de tiempo (Ventolo y Williams, 2012). El valor del suelo, por lo tanto, se refiere a la corriente de ingresos generados por el suelo como un factor de producción (una vez expuesto a mejoras) situado en un período determinado de tiempo. En este caso el valor está vinculado a la ventaja de la generación de ingresos que provee el suelo. Sin embargo, incluso el suelo que es colocado en el proceso de generación de ingresos, de igual manera generará un valor derivado apoyándose en el concepto del costo de oportunidad (Kauko, 2003). El costo de oportunidad está sujeto a las condiciones prevalecientes del mercado. Como factor de producción, el suelo al igual que cualquier otro producto, se comercializa en el mercado. Esto significa que el valor del suelo será objeto de la oferta y la demanda. A este respecto, el suelo se convierte en un producto único, ya que cuenta con una oferta y demanda única (Appraisal Institute, 2013).

Al momento de considerar la valoración del suelo se debe comprender que el mismo es fijo en términos de ubicación geográfica. Al respecto, McCluskey (2005) indica que, por tal razón, todas las ventajas económicas proporcionadas por el suelo deben ser utilizadas en el lugar donde ubica éste. Se debe concluir que la localización es crucial para determinar el valor del suelo debido a que la escasez de la oferta en un lugar no puede ser sustituida por excedentes en otro. En esta línea, al estudiar los procesos de valoración de inmuebles, Orford (1999) plantea que el valor del suelo se verá influido por los factores económicos relacionados con su localización y a su vez la inmovilidad del suelo influirá en sus

características económicas. McCluskey (2005), plantea que al igual que el suelo se encuentra en una ubicación geográfica fija, su oferta también se considera como fija. Añade a este planteamiento que, dado que la oferta por el suelo es fija, en teoría, su curva de oferta es perfectamente inelástica, lo que significa que los aumentos en el precio del suelo derivados de un cambio en la curva de demanda, no estimulan un aumento de la oferta. La oferta seguirá siendo la misma sin importar el incremento en los precios.

Tomando en consideración estos aspectos, se puede concluir que el suelo no es demandado por sus propios méritos, sino por su capacidad de producción de un bien o un servicio por medio de la mejora que sufra. Por ejemplo, los aumentos en la demanda de vivienda ocasionan una mayor demanda de suelo para producción residencial. Dado que el uso del suelo es intercambiable, esto afectará la oferta para ese uso y por tanto el valor. La demanda por otros bienes y servicios y los ingresos procedentes de estos puede determinar los ingresos devengados por el suelo para un uso determinado y, por ende, su valor.

Al plantear el concepto valor y relacionarlo con el suelo, se pueden discutir diferentes acepciones de esta relación. Entre estas se encuentra el valor de mercado, valor de utilización y valor catastral. Es este último el de mayor importancia para la presente investigación. Define el Appraisal Institute de los Estados Unidos que por valor catastral se debe entender como *“aquel que rige en la tributación ad valorem y hace referencia al valor de una propiedad de acuerdo con el registro de contribuyentes de un territorio”* (Appraisal Institute, 2013). Se puede dar la situación que este tipo de valor no concuerde con el valor de mercado, aunque se entiende que su cálculo está relacionado con éste.

Los pasados años se ha experimentado una situación económica precaria en muchos países del mundo. En los Estados Unidos la burbuja económica generada por el mercado inmobiliario estalló y colocó la economía en general en una situación precaria (Coy, 2008; Krugman, 2009). Aunque se puede constatar una mejoría de esa situación, no se puede asegurar que se ha superado totalmente (Krugman, 2014). Al estudiar la dinámica de valores inmobiliarios se puede establecer que son cuatro los factores que influyen en su valor (Rattermann, 2004; Hong y Brubaker, 2010):

- Tendencias sociales como pueden ser cambios poblacionales
- Controles del sector gubernamental, por ejemplo, calificación del suelo o políticas fiscales

- Condiciones ambientales, tales como zonas contaminadas
- Circunstancias económicas tales como evolución del empleo o de oferta y demanda

Este último es fundamental en la condición del mercado inmobiliario actual.

Cuando se revisan los conceptos teóricos que establecen la base para el entendimiento del comportamiento de los mercados inmobiliarios (Rattermann, 2004; Hong y Brubaker, 2010; Jacobus, 2010; Shiller, 2011; Ventolo y Williams, 2012) se puede desprender que el mercado inmobiliario genera una gran demanda por suelo. Para que este mercado opere de manera efectiva es necesario que en el proceso de oferta y demanda se materialice la transacción comercial.

El funcionamiento del mercado inmobiliario en las ciudades está íntimamente relacionado a la dinámica de crecimiento espacial y del crecimiento económico urbano. En la situación económica reciente en Puerto Rico se puede evidenciar como el mercado inmobiliario ha generado una oferta basada en procesos de financiamiento hipotecario alterados y valores residenciales inflados, los que al enfrentar la realidad han hecho que muchos individuos pierdan sus residencias y los que desean adquirir no cuenten con las credenciales financieras necesarias para adquirir un inmueble (Shiller, 2011). Esto a su vez va creando un mercado donde la oferta domina y la demanda se ve menguada por falta de liquidez económica y timidez a la inversión.

El problema estriba en que el suelo es la materia primaria en el proceso de crecimiento urbano (Jacobus, 2010). Los intereses económicos de sus propietarios entran en conflicto de forma casi permanente con los promotores inmobiliarios. Como el suelo urbano es un bien escaso, que no sufre deterioro con el paso del tiempo puede alcanzar rápidas revalorizaciones y por tanto es susceptible a maniobras especulativas por parte de sus propietarios. Al estudiar este proceso se observa cómo, en muchas ocasiones, los propietarios conservan el suelo con potencial de crecimiento urbano a la espera que el valor de éste aumente debido la demanda. Se debe entender que el concepto de escasez del suelo urbano va en función a la capacidad de desarrollo del mismo (Rattermann, 2004). Por ejemplo, los suelos rústicos que no sean aptos para el desarrollo no entran en el proceso de demanda. Igualmente, los suelos rústicos que puedan ser urbanizados pueden llegar a no ser suficientes para generar la oferta requerida. Por tanto, se debe entender como escasez en el mercado y no en términos absolutos.

Otro aspecto fundamental es el de los desequilibrios que se producen como consecuencia de la falta de elasticidad de la oferta ante el cambio de la demanda. Es evidente que la demanda poblacional ejerce una gran influencia. Sin embargo, el proceso de adecuar suelo para ser ocupado por la ciudad es uno que consume largos plazos de tiempo, particularmente en la fase de planeamiento (Rattermann, 2004).

También las condiciones geomorfométricas del suelo ejercen una influencia directa sobre el valor de éste. Los suelos que presenten mayor dificultad de desarrollo basado en estos aspectos tendrán un valor más bajo de mercado (Ferrando, 2004). Por ejemplo, si el suelo es modificado para acomodar el proceso de urbanización (alteración topográfica, canalización cuerpos de agua) el costo de mejora es trasladado al inmueble construido. Por otra parte, las estipulaciones legales establecidas en el proceso de planeamiento definen el uso e intensidad con que se puede urbanizar el suelo (Appraisal Institute, 2013). Dado que el valor del suelo surge como una razón de la capacidad de desarrollo con que cuente el mismo, se debe entender que cuando mayor sea la intensidad de uso, mayor será el valor del suelo. Se puede entender entonces que el precio final de las residencias o la estructura que se construyan incluye aquel establecido para el suelo.

En la fórmula sobre el valor del suelo se debe considerar también el efecto de la localización (McCluskey, 2005). Históricamente los geógrafos y los economistas han estado investigando las relaciones existentes entre medios de transporte, patrón de ocupación del suelo, el valor del suelo y las propiedades. Los trabajos de von Thunen, Alonso y Christaller son evidencia del trabajo que sirve como base teórica para las investigaciones contemporáneas. En estos se resume que las posiciones centrales de la ciudad son preferidas por el ahorro que suponen en desplazamientos. También se puede dar una sobrevalorización en razón de la calidad urbanística medida en términos de urbanización, equipamientos, estructura social y otros. Tomando esta idea como principio se puede concluir que a más lejos del centro y sus atractivos, más barato será el suelo. En un sistema de libre mercado el que más pueda pagar ocupará los mejores suelos.

2.2 Fundamentos de la valoración de inmuebles

La valoración o tasación de inmuebles es una opinión sobre el valor del inmueble (Appraisal Institute, 2013). Esta opinión debe estar apoyada por datos generales y específicos, debe ser a

una fecha específica, el valor estimado debe ser definido, el inmueble que se está valorando debe estar claramente descrito y la persona que realiza la valoración debe ser totalmente imparcial y no tener ninguna relación o interés con el inmueble (Ventolo y Williams, 2012).

Esta estimación se le conoce como valor de mercado. Definido propiamente el valor de mercado es la cantidad estimada por la cual, en la fecha de la valoración, se intercambiaría voluntariamente una propiedad entre un comprador y un vendedor en una transacción libre después de una comercialización adecuada en la que cada una de las partes ha actuado prudentemente y sin presiones (Appraisal Institute, 2013).

No se debe confundir esto con el precio del inmueble. Por precio se entiende la cantidad de dinero que un comprador particular está realmente dispuesto a pagar y un vendedor particular está dispuesto a recibir en las circunstancias concretas de una transacción.

Al momento de tasar un inmueble es importante tener presente tres principios (Appraisal Institute, 2013; Ventolo y Williams, 2012): principio de sustitución, principio de contribución y principio de mayor y mejor uso.

Por principio de sustitución se entiende que el valor de un inmueble es equivalente al de otros de similares características sustitutivas de aquel igualmente deseables, que dan igual rendimiento o prestan los mismos servicios o utilidad y de valor equivalente. Bajo el principio de contribución (o de valor residual) se entiende que un componente de la propiedad depende de su contribución al valor total de ella. Por último, el principio de mayor y mejor uso establece que el valor de un inmueble, susceptible de ser dedicado a diferentes usos o de ser construido con distintas intensidades edificatorias, es el que resulta de ser destinado, dentro de las posibilidades legales y físicas, al uso más razonable, probable y financieramente aconsejable que le permita obtener el mayor valor.

2.3 El Proceso de Valoración

El proceso de valoración surge de la necesidad de asignar valor a un inmueble en un proceso de compra-venta; en ese momento se busca una estimación imparcial del valor de mercado del mismo (Rattermann, 2004).

Esta valoración requiere de un proceso cuya aplicación va en función del carácter de la valoración (Tabla 2.1, Appraisal Institute, 2013). Los pasos de este proceso toman como

inicio la definición del problema. Con esta acción se fijan los límites de la valoración y se elimina toda ambigüedad que pueda surgir en el proceso. La identificación se fundamenta en los aspectos espacio-temporales del inmueble como pueden ser el lugar donde ubica el mismo, el área que ocupa y la fecha de la estimación. A estos aspectos se suma la identificación de derechos de propiedad que serán valorados y el objetivo de la valoración (compra, venta, tributación, expropiación).

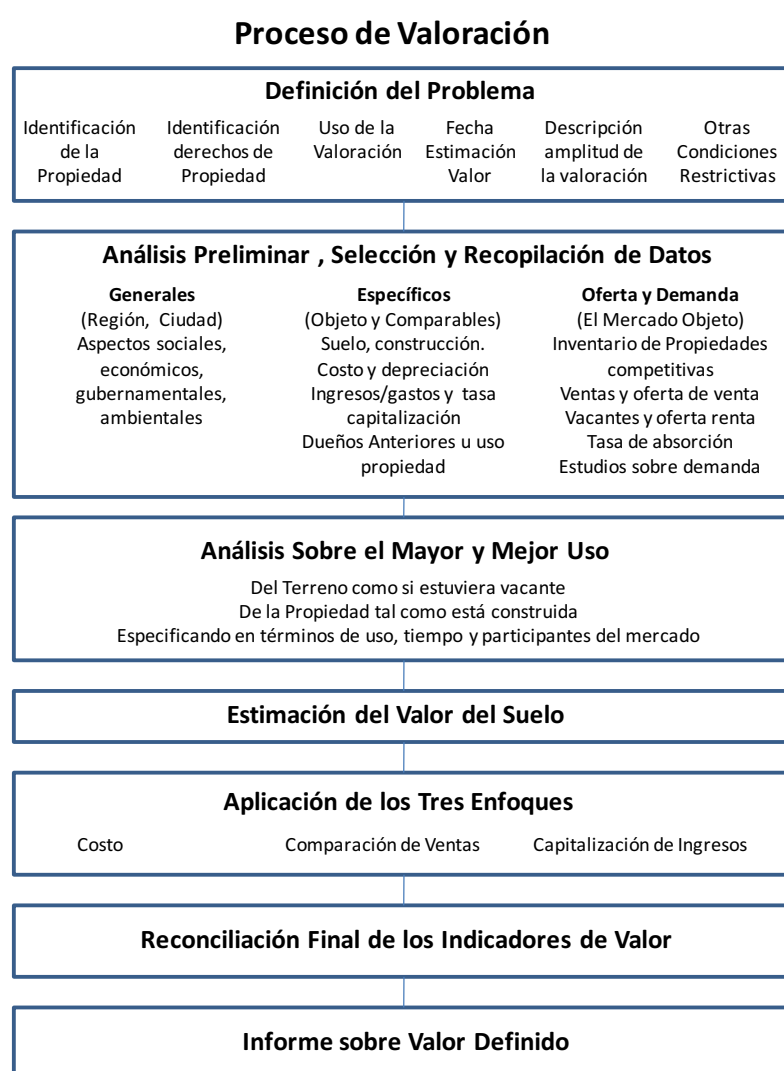


Tabla 2.1: Proceso de Valoración (Adaptado de "The Appraisal of Real Estate" Edición 14. Appraisal Institute, 2013.

Una vez se define el problema se procede a seleccionar y recopilar datos para el análisis del mismo. Los datos se recopilarán en función del tipo de inmueble a valorar. En general se recopilan datos en tres aspectos: datos generales, datos específicos y datos sobre oferta y

demanda. Los datos generales buscan establecer las tendencias en las fuerzas sociales, económicas, gubernamentales y ambientales que pueden afectar el valor del inmueble. A modo de ejemplo, en la actualidad la situación económica ha generado una baja en la compra de residencias. Al sumar la cantidad de residencias construidas en los años recientes que no han sido compradas, más las residencias que se ofertan en venta debido a la situación precaria de los dueños, resulta un exceso en la oferta.

Los datos específicos se relacionan con el inmueble objeto de avalúo y propiedades comparables. Se incluyen datos como la ubicación, aspectos físicos, costos, gastos, ingresos de la propiedad. Por otro lado, los datos de oferta y demanda se relacionan con la posición competitiva del inmueble en un mercado futuro. Se consideran inventarios de propiedades, tasas de ocupación, población.

2.3.1 Enfoques de Valoración de Inmuebles

Existen varios enfoques que son utilizados al momento de establecer el valor de un inmueble (Ballester y Rodríguez, 1999; Ferrando, 2008; Ventolo y Williams, 2012). Entre estos se puede mencionar el enfoque de valor residual, el enfoque por costo de reposición, la capitalización de rendimientos y, el de mayor uso, el enfoque por comparación del mercado. Es éste último el que se utiliza en Puerto Rico para determinar el valor de los inmuebles al momento de realizar un proceso de venta. Por lo tanto, tiene una gran influencia al momento de acordar un precio entre las partes. Por tal razón toma relevancia conocer con mayor detalle el mismo.

El enfoque por comparación de mercado permite determinar el valor de la propiedad que se tasa según la comparación de valores de ventas de propiedades similares captadas del mercado inmobiliario (Aravena, 2007). El enfoque busca reflejar la naturaleza del bien, las condiciones de mercado del momento y las circunstancias más probables bajo las cuales ese bien se comercializaría en el mercado. Además, se busca reproducir los criterios y procesos deductivos de los participantes en ese mercado. Para este enfoque es fundamental determinar el valor a partir de información específica y concreta del mercado apropiado. Cuenta, sin embargo, con una serie de limitaciones (Aravena, 2007).

- Supone un mercado con competencia perfecta.
- No todos los inmuebles tienen un mercado amplio y representativo.

- Requiere información suficiente (cantidad y calidad).
- La información es variable en el tiempo.

La ecuación para resolver la tasación de un inmueble según el enfoque de comparación de mercado es la siguiente (Aravena, 2007):

$$Ti = \text{Sup m}^2 * UF/M^2 \text{ (ESM)}$$

dónde:

- Ti = Valor de tasación del inmueble
- Sup m^2 = Magnitud de superficie edificada o de suelo en el caso de la tasación de terrenos.
- UF/M^2 = Valor en unidad de medida de lo edificado o de suelo, en el caso de terrenos. Este valor incluye el valor del suelo, de lo edificado y de las obras complementarias. En los inmuebles de copropiedad, incluye la cuota proporcional del terreno, de las áreas comunes y de las obras complementarias.⁴
- ESM = Estadística sistematizada del mercado la cual permite establecer una franja de valores (varianza) con promedios, máximos y mínimos.

Al utilizar el enfoque por comparación se realizan una serie de pasos que ayudan a recopilar y analizar la propiedad estableciendo una serie de características y cualidades que influyen en el valor de la propiedad (Appraisal Institute, 2013; Aravena, 2007). Estos son:

- a. Estudio de mercado inmobiliario: Se contempla el análisis de la propiedad en relación a su mercado basado en estadísticas sistematizadas de mercado. Esto implica la constatación de la existencia de un mercado representativo de inmuebles comparables, es decir, de atributos similares. También se debe corroborar la existencia de información actualizada y confiable del comportamiento de dicho mercado.
- b. Recopilación de información: La recopilación de la información de mercado se basa en recopilar datos de oferta y/o de demanda del mercado inmobiliario. Es importante considerar la presencia de variables espaciales y temporales. En términos estadísticos, a mayor número de referencias, mayor exactitud y fiabilidad de la información; a mayor

⁴ La UF/M^2 del suelo en área urbana caracterizada por un alto coeficiente de constructibilidad (uso de suelo en alta densidad y edificación en altura) incluye habitualmente el valor UF/M^2 de la edificación que allí exista.

información específica, mejores resultados de los ajustes; a mayor semejanza, más significativa es la información. En el caso de escasez de información es importante reconocer el hecho y siempre tener presente que la muestra no es aleatoria. Los bienes son heterogéneos y por lo tanto son bienes sustitutos, similares, pero no idénticos.

Los datos a seleccionar deben contar con una serie de consideraciones que fortalecen la opinión de valor final. Primero, una clasificación de las tipologías de inmuebles según su uso y tipos edificatorios (vivienda aislada, bloque, edificio, etc.). Segundo, una definición clara de las características del inmueble que se desea valorar para así conseguir inmuebles con características similares en el mercado (igual cantidad de dormitorios, similitud en área de cocina o cantidad de baños). También es importante ser preciso en las características de la localización de los inmuebles, la homogeneidad del sector, densidad del sector o estrato socioeconómico del sector. Finalmente, la fecha del valor de venta de los inmuebles seleccionados para la comparación.

- c. Verificación de la información o control de calidad: La información recopilada requiere un proceso de verificación. Este proceso busca descartar o validar datos. Entre las variables para tomar alguna de estas acciones se encuentra la existencia de datos incompletos o repetidos. Se busca identificar heterogeneidad de construcciones.
- d. Análisis de la propiedad en su mercado: El análisis de la propiedad a tasar en relación a su mercado implica su comparación con otras que compiten por igual demanda. Por ello el análisis comparativo debe captar inmuebles similares en ubicación, características del entorno inmediato y accesibilidad, características del terreno, condición de la construcción, análisis del mayor y mejor uso.
- e. Reconciliación de valores: A través de un procedimiento estadístico se establece una reconciliación entre los valores de las referencias captadas más similares al del inmueble tasado. Con esto se establece una aproximación matemática del valor de tasación ya que se determina un valor promedio con la correspondiente desviación estándar. También se logra un margen de error de la estimación deducido de los datos de la “franja de valor”. Obteniendo el valor promedio señalado se puede proceder a resolver la ecuación de tasación por comparación de mercado.

2.3.2 Valoración Hedónica

La técnica de valoración hedónica se fundamenta en la idea que si un bien está constituido por una serie de atributos, entonces se puede asumir que el precio del mercado del bien surge de la suma de los precios individuales de los atributos (Figueroa y Lever, 1992; Monson, 2009). Visto de otra manera, la teoría de valoración hedónica permite identificar la importancia relativa de cada atributo en el valor asignado por el mercado a un bien, mediante lo cual es posible determinar cómo cambiará dicho valor al variar la cantidad y calidad en que se encuentra presente cada uno de estos atributos y consecuentemente predecir precios (Figueroa y Lever, 1992; Caridad et al, 2008).

La metodología que se utiliza para obtener el valor se basa en la utilización de técnicas estadísticas de regresión que ayuden a describir la relación entre el precio del bien y sus respectivas características (Dorsey, 2010). En el caso de los bienes inmuebles, la teoría pretende explicar el valor del inmueble considerando dentro del grupo de características condiciones tales como; superficie, uso del suelo, calidad de la construcción, ubicación, vecindario y otros (Liu et al, 2010). En general éstas se pueden agrupar en dos categorías; las relacionadas con la condición estructural del inmueble y las relacionadas con la localización y su entorno.

La primera aplicación de las teorías hedónicas a los bienes inmuebles fue el estudio de Ridker y Henning (1967) donde se analizó el efecto de la contaminación de aire sobre el valor de las viviendas en la ciudad de St. Louis en Estados Unidos. Posterior a Ridker y Henning se encuentra el trabajo de Rosen (1974) quien formuló un modelo hedónico de dos etapas donde primero se obtienen los precios del bien y segundo las demandas implícitas para cada atributo del bien.

A partir de la aportación de Rosen (1974), se generó una gran cantidad de investigaciones utilizando el modelo hedónico en la valoración de inmuebles residenciales. Esto queda evidenciado al realizar una revisión a la literatura relacionada al tema. En la década del 1980 y 1990, la mayoría de estas se aplicaron en ciudades de los Estados Unidos. Luego a partir del año 2000 se comienza a observar mayor uso de este modelo en ciudades europeas y asiáticas. Dentro de esos estudios iniciales se puede destacar los trabajos de Li y Brown (1980), Gillard (1981), Walden (1990), Sirpal (1994) y el de Rodríguez y Sirmans (1994). Junto al de Rosen (1974), estos han servido de referencia a una gran cantidad de estudios realizados en las últimas dos décadas. Todos estos, aunque enfatizados en la valoración del

inmueble, buscaban entender el desempeño particular de alguna variable dentro de todas las seleccionadas para el análisis. Por ejemplo, Gillard (1981) buscaba establecer la importancia de amenidades ambientales en el valor del inmueble residencial o Sirpal (1994), que buscaba comprender la influencia de la cercanía a una escuela en el valor de inmuebles residenciales. En la actualidad esa tendencia se mantiene. Las variables que se busca estudiar se pueden organizar en aquellas que buscan la relación de diversos aspectos ambientales, los que atienden temas de infraestructura y un tercer grupo de temas generales. En cuanto a los temas de tipo ambiental se puede mencionar los trabajos de Anselin y Gallo (2006) que estudian la influencia de la calidad de aire en el valor inmueble, Zygmunt y Gluszak (2015) que estudian cuanto influye en el valor inmueble la cercanía a zonas boscosas y Fu et al. (2016) que buscan definir el impacto en el valor de residencias del cambio climático, en particular el aumento en el nivel del mar en las costas.

En cuanto a temas de infraestructura predominan los relacionados a los sistemas de transporte. De estos se puede destacar los trabajos de Du y Mulley (2006), Dube et al. (2013), Sun et al (2016) y Zhong y Li (2016). Todos estos son trabajos dedicados a comprender el efecto que tienen los sistemas de transporte de trenes en ciudades en el valor de los inmuebles. Similar a estos estudios se encuentra el de Mulley y Tsai (2017) cuya diferencia estriba en el análisis de sistemas de autobuses.

El modelo hedónico ha sido aplicado utilizando diversos procesos matemáticos y estadísticos. El más común en los estudios iniciales fue la técnica de los Mínimos Cuadrados Ordinarios (MCO). En años recientes se observa una tendencia creciente de estudios donde se compara esta con otros métodos o simplemente se utilizan estos en sustitución del MCO. Dentro de estos destaca el uso de redes neuronales. Ejemplos de estos son la valoración en masa realizada por Peterson y Flanagan (2009) o el estudio de Kutasi y Badics (2016) aplicado en Budapest.

En ésta línea, la Regresión Geográficamente Ponderada también ha sido utilizada. Un trabajo seminal con esta línea de investigación es el de Brunsdon et al. (1999). De gran importancia es el estudio de Bitter et al. (2007) en el que los autores buscaron determinar la heterogeneidad espacial de valores residenciales al comparar el método de la Regresión Geográficamente Ponderada contra el Método de Expansión Espacial con datos de la ciudad de Tucson en Arizona. Se debe destacar el trabajo de Yu (2007) y el de Lu et al. (2014).

Como se ha discutido anteriormente, el aspecto temporal del valor inmueble es uno de los ejes principales de la presente investigación. Por tal razón este tema también fue tomado en consideración al realizar una revisión a la literatura existente. Sobre este tema se pudo identificar una reducida cantidad de estudios. Sin embargo, ninguno de ellos discutía el aspecto de la legitimidad en el tiempo de un valor inmueble para aplicaciones de valoración catastral. En su mayoría los estudios analizan los cambios en el tiempo de los valores inmuebles utilizando la técnica de valoración hedónica. Ejemplo de estos estudios son el de Dubé y Legros (2014), Swoboda et al (2015) o Devaux y Dubé (2016).

La revisión a la literatura también permitió identificar estudios donde la singularidad de la variable que se buscaba estudiar no puede clasificarse dentro de los grupos antes mencionados. Ejemplos de estos son; el estudio del impacto de sistemas fotovoltaicos en el valor de inmuebles realizado por Wee (2016), la influencia que ejercen las facilidades deportivas en el valor de residencias realizado por Feng y Humphreys (2016) y el estudio en que se relaciona la existencia de ofensores sexuales en una comunidad y el impacto en el valor de las residencias.

Capítulo 3

Modelación de Valores de Ventas de Inmuebles Residenciales en un Entorno SIG

3. Modelación de Valores de Venta de Inmuebles Residenciales en un Entorno SIG.

Ya analizados y caracterizados los elementos configuradores del problema de valoración de inmuebles con fines contributivos, se cuenta con los elementos suficientes para definir un modelo que ayude a mejorar el mismo. Se busca un método que permita de manera ágil establecer los valores de inmuebles en una comunidad. Puerto Rico muestra la particularidad que, a partir de la década de 1950, la inmensa mayoría de los desarrollos urbanos se basan en el modelo de urbanización en el cual todas las residencias del complejo urbanístico son idénticas en términos de tipología edificatoria. Al establecer el valor de un inmueble con algún método adoptado, se puede asumir que todas las residencias dentro de la urbanización varían muy poco en su valor. Este principio es el que se toma de base para la tasación por métodos comparables. Al contar con esta condición se puede adoptar un modelo que, fundamentado en datos conocidos, establezca un valor del inmueble para la urbanización donde se localiza. Los datos conocidos para ejecutar el modelo surgen de las ventas de inmuebles ocurridas en el año anterior para el cual se desea adoptar un nuevo valor. Al contar de manera constante con valores de nuevas ventas realizadas, el modelo se puede actualizar anualmente.

Para la predicción del valor residencial en la comunidad se fundamentará en modelos estadísticos. Para esto los más utilizados son la técnica de valoración hedónica fundamentada en los mínimos cuadrados ordinarios o la técnica de Regresión Geográficamente Ponderada. Con un mismo grupo de datos se procederá a aplicar estos métodos y se comprobará su certeza en la predicción de los valores comparando estos con valores reales de ventas. El método cuya predicción sea más cercana a las ventas reales será el adoptado para la aplicación final.

A continuación, se expone una discusión sobre la técnica de valoración hedónica fundamentada en los mínimos cuadrados ordinarios y la técnica de Regresión Geográficamente Ponderada.

3.1 Valoración Hedónica en un entorno SIG

El objetivo de la valoración hedónica es establecer el valor de una variable dependiente, que en este caso es el valor del inmueble, basándose en una serie de características del mismo (variables independientes o explicativas). La identificación de éstas variables independientes o explicativas va en función de la hipótesis que se formula sobre la influencia de los diversos factores que conforman el valor buscado. En ocasiones la selección y utilización de estas variables depende del acceso a los datos. Sobre la selección de éstas, Caridad et al. (2008) plantean que es deseable incluir un número no demasiado elevado de este tipo de variables por dos razones principales; primero, porque las variables independientes suelen estar relacionadas entre sí y, en segundo lugar, porque la inclusión de un número elevado de variables no origina importantes mejoras en el poder explicativo del modelo.

3.2 Mínimos Cuadrados Ordinarios

De todas las técnicas de regresión utilizadas para el modelo de valoración hedónica la más conocida es la de los Mínimos Cuadrados Ordinarios (Yaolin, 2006; Kestens et al., 2006; Liu et al., 2010; Dorsey et al., 2010; Seo et al., 2014; Morano y Tajani, 2014; Klovers y Pereira, 2015). La técnica de los Mínimos Cuadrados Ordinarios (MCO) se considera un punto de inicio adecuado para todos los análisis de regresión espacial. Esta técnica proporciona un modelo global de la variable que intenta predecir ya que crea una ecuación de regresión simple para representar ese proceso.

Al revisar literatura relacionada (Yaolin, 2006; Kestens et al., 2006; Liu et al., 2010; Dorsey et al., 2010; Seo et al., 2014; Morano y Tajani, 2014; Klovers y Pereira, 2015) se destacan dos razones para utilizar el análisis de regresión:

- Entender el grado de relación entre una variable de estudio y otras variables que ayudan a definir ésta (grado de relación entre variable dependiente e independientes).
- Predecir valores en otras localizaciones o tiempos.

El análisis de regresión se efectúa por medio de la ecuación de regresión. Esta es la fórmula matemática que se aplica a las variables explicativas para predecir mejor la variable dependiente que intenta modelar. Cada variable independiente se asocia con un coeficiente de

regresión que describe la fortaleza y el signo de la relación de esa variable con la variable dependiente. La forma general de la ecuación de regresión múltiple se muestra a continuación:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \dots \beta_n X_n + \epsilon$$

En esta la y es la variable dependiente, las X son las variables explicativas y las β son coeficientes de regresión. Estos componentes se explican en más detalle a continuación:

- Variable dependiente (y): Representa el proceso que intenta prever o entender.
- Variables independientes o explicativas (X): Variables que se utilizan para modelar o prever los valores de la variable dependiente. En la ecuación de regresión, aparecen en el lado derecho del signo igual y a veces se denominan variables explicativas. La variable dependiente es una función de las variables explicativas.
- Coeficientes de regresión (β): Los coeficientes se calculan con la herramienta de regresión. Son valores, uno para cada variable explicativa, que representan la fortaleza y el tipo de relación que tiene la variable explicativa con la variable dependiente. Cuando la relación es positiva, el signo del coeficiente asociado también es positivo. Los coeficientes de las relaciones negativas tienen signos negativos. Cuando la relación es una relación sólida, el coeficiente es relativamente grande (relativo a las unidades de la variable explicativa con la que se asocia). Las relaciones débiles se asocian con coeficientes cercanos a cero. Representa el valor esperado para la variable dependiente si todas las variables independientes (explicativas) son cero.

La creación de un modelo de regresión es un proceso iterativo que implica buscar variables independientes efectivas para explicar la variable dependiente que intenta modelar o entender, ejecutar la herramienta de regresión para determinar qué variables son indicadores efectivos, quitar o agregar repetidamente variables hasta que encuentre el mejor modelo de regresión posible (Liu, 2010). El proceso de creación del modelo debe identificar posibles variables explicativas al consultar la teoría, expertos en el campo y el sentido común. Debe poder establecer y justificar la relación esperada entre cada variable explicativa posible y la variable dependiente antes del análisis, y debe cuestionar los modelos donde estas relaciones no coinciden (Yaolin, 2006).

La regresión de MCO es un método simple, tiene una teoría bien desarrollada y posee una cantidad de diagnósticos efectivos para ayudar con la interpretación y la solución de problemas. Sin embargo, MCO sólo es efectivo y confiable si los datos y el modelo de regresión satisfacen todas las presuposiciones requeridas intrínsecamente por este método (Rosenshein et al., 2011). Los datos espaciales a veces violan las presuposiciones y requisitos de la regresión de MCO, por eso es importante utilizar las herramientas de regresión en conjunto con herramientas de diagnóstico apropiadas que puedan evaluar si la regresión es un método apropiado para el análisis, dada la estructura de los datos seleccionados.

El mayor error en la aplicación de modelos de regresión radica en su diseño (Rosenshein et al., 2011). Un modelo mal diseñado es el que no está completo, le faltan variables explicativas importantes, por eso no representa en forma adecuada lo que está tratando de modelar o predecir. Un mal diseño es evidente siempre que vea una autocorrelación espacial estadísticamente significativa en los residuales de regresión o, dicho de otra manera, siempre que observe que las predicciones excesivas y escasas (residuales) de su modelo tienden a agruparse espacialmente.

El modelo hedónico ha sido desarrollado históricamente con presuposiciones que no contemplan el factor espacial (Brunsdon et al., 1998). Sin embargo, la integración de estas herramientas en los SIG hace cada día más frecuente la utilización de estos modelos desde la óptica espacial. Además, esto ha permitido la creación de nuevos modelos donde el factor espacial toma un papel importante. Los datos espaciales exhiben dos propiedades que hacen difícil, pero no imposible, cumplir con las presuposiciones y los requisitos de los métodos estadísticos no espaciales (Rosenshein et al., 2011).

- En general, las entidades geográficas están autocorrelacionadas espacialmente; esto significa que las entidades cercanas tienden a ser más similares que las entidades que más alejadas entre sí. Esto crea un tipo de influencia que cuenta más para los métodos de regresión (no espaciales) tradicionales.
- La geografía es importante y a veces los procesos más importantes para lo que se está modelando no son estacionarios, sino que se comportan de manera diferente en

distintas partes del área de estudio. Esta característica de los datos espaciales se puede denominar variación regional o no estacionariedad.

Los verdaderos métodos de regresión espacial se desarrollaron para manejar estas dos características de los datos espaciales e incluso incorporar estas cualidades particulares de los datos espaciales para mejorar su capacidad de modelar las relaciones de los datos. Algunos métodos de regresión espacial tratan de manera efectiva con la primera característica (autocorrelación espacial), otros tratan de manera efectiva con la segunda (no estacionariedad).

Un estadístico que aplica estos modelos de la manera tradicional ve el efecto de la autocorrelación espacial muy diferente a como lo ve un estadístico espacial. El estadístico tradicional la ve como un factor negativo que se debe eliminar de los datos a través del remuestreo porque la autocorrelación espacial viola las presuposiciones subyacentes de varios métodos estadísticos no espaciales tradicionales (Tu et al., 2007; Smith, et al., 2015). Sin embargo, para el geógrafo o analista de SIG, la autocorrelación espacial es evidencia de procesos espaciales subyacentes importantes en el trabajo; es un componente integral de los datos (Liu et al., 2010). Al eliminar el factor espacio, se extraen los datos de su contexto espacial. Los procesos espaciales y las relaciones espaciales demostradas en los datos son de interés principal y una de las razones del crecimiento de los SIG. Sin embargo, para evitar un tipo de influencia que cuente más en su modelo, se debe identificar el conjunto completo de variables explicativas que efectivamente capturen la estructura espacial inherente en la variable dependiente (Liu et al., 2010). Si no se puede identificar todas estas variables, probablemente se observará la autocorrelación espacial estadísticamente significativa en los residuales del modelo. Lamentablemente, los resultados de regresión no serán confiables hasta que esto se solucione.

Para tratar con la autocorrelación espacial en los residuales del modelo de regresión Rosenshein, et al. (2011) identifican tres estrategias que se pueden seguir:

- a. Re-muestrear hasta que las variables de entrada ya no exhiban una autocorrelación espacial estadísticamente significativa. Aunque esto no garantiza que el análisis está libre de problemas de autocorrelación espacial, son mucho menos probables cuando se reduce la autocorrelación espacial de las variables dependientes y explicativas. Este

es el enfoque del estadístico tradicional para tratar con la autocorrelación espacial y sólo es apropiado si la autocorrelación espacial es el resultado de la redundancia de los datos.

- b. Aislar los componentes espaciales y no espaciales de cada variable de entrada mediante un método de regresión de filtrado espacial. El espacio se quita de cada variable, pero se vuelve a colocar en el modelo de regresión como una nueva variable para dar cuenta de los efectos espaciales/la estructura espacial.
- c. Incorporar la autocorrelación espacial en el modelo de regresión mediante métodos de regresión econométricos espaciales.

Los modelos globales, como son los MCO, crean ecuaciones que describen mejor las relaciones de datos generales en un área de estudio. Cuando esas relaciones son consistentes en el área de estudio, la ecuación de regresión de MCO modela bien esas relaciones (Burt et al., 2009). Sin embargo, cuando esas relaciones se comportan de manera diferente en distintas regiones del área de estudio, la ecuación de regresión es un promedio de la mezcla de relaciones presentes. En el caso donde esas relaciones representan dos extremos, el promedio global no modelará bien ningún extremo. Cuando las variables explicativas exhiben relaciones no estacionarias (variación regional), los modelos globales tienden a apartarse a menos que se utilicen métodos sólidos para calcular los resultados de regresión (Rosenshein et al., 2011). Idealmente, podrá identificar un conjunto completo de variables explicativas para capturar la variación regional inherente en su variable dependiente. Sin embargo, si no puede identificar todas estas variables espaciales, volverá a observar una autocorrelación espacial estadísticamente significativa en los residuales del modelo o valores R^2 menores de lo esperado. Los resultados de regresión no serán confiables hasta que esto se solucione.

De las teorías relacionadas, se puede enumerar al menos cuatro formas de tratar con la variación regional en los modelos de regresión de MCO (Rosenshein et al., 2011; Smith et al., 2015; Brunsdon et al., 1998)

- a. Incluir una variable en el modelo que explique la variación regional.
- b. Utilizar métodos que incorporen variación regional en el modelo de regresión como el de la Regresión Ponderada Geográficamente.
- c. Consultar las probabilidades y los errores estándar de regresión sólidos para determinar si los coeficientes de la variable son estadísticamente significativos.

- d. Redefinir o reducir el tamaño del área de estudio de manera que los procesos dentro de ésta sean todos estacionarios (para que ya no exhiban una variación regional).

Al aplicar las técnicas de regresión, la tarea más importante es la selección acertada de variables explicativas que permitan definir el modelo de manera acertada. Es importante contar con una serie de pruebas diagnósticas que permitan seleccionar de la manera más efectiva estas variables. Una vez identificada la variable dependiente y seleccionadas las variables explicativas se pasa a la aplicación de la técnica de regresión seleccionada. En este caso se utilizará la de los Mínimos Cuadrados Ordinarios (MCO). Ésta genera una serie de pruebas diagnósticas que permiten establecer la calidad del modelo. El programa de SIG ArcGIS es el seleccionado para realizar el análisis. En el caso de este programa la aplicación de los MCO genera dos productos que permiten interpretar los resultados. El primero es el informe de resumen de MCO. El segundo resultado es un mapa de residuales de regresión.

Uno de los aspectos principales a analizar es el nivel de apoyo de las variables explicativas independientes al modelo en general (Rosenshein et al., 2011). Para determinar qué variables son más significativas la herramienta de MCO en ArcGIS calcula el coeficiente para cada variable explicativa y luego ejecuta una prueba estadística para determinar si la variable ayuda al modelo o no. Esta prueba estadística calcula la probabilidad que el coeficiente sea cero. Si en esta prueba la variable explicativa obtiene un cero o un valor cercano a cero, significa que la misma representa un impacto muy bajo para el modelo. Mientras más baja sea la probabilidad, más baja es la probabilidad que el coeficiente sea cero. En el caso del programa ArcGIS, cuando la probabilidad es menor a 0,05, un asterisco próximo a la probabilidad en el informe del MCO indica que la variable explicativa ayuda al modelo, su coeficiente es estadísticamente significativo al 95% de confianza. Si se buscan variables explicativas estadísticamente significativas se debe buscar aquellas que se muestran en el informe con un asterisco.

Otro aspecto a observar es la estacionariedad del modelo (Rosenshein et al., 2011). Para determinar ésta, se calcula la prueba estadística de Koenker (Breusch-Pagan estudentizado de Koenker) para la no estacionariedad. Esta es una prueba que ayuda a determinar si las variables explicativas del modelo tienen una relación consistente con la variable dependiente, tanto en el espacio geográfico como en el espacio de datos. Cuando el modelo es consistente en el espacio geográfico, los procesos espaciales representados por las variables explicativas

se comportan de la misma manera en cualquier parte del área de estudio, los procesos son estacionarios. Cuando el modelo es consistente en el espacio de datos, la variación en la relación entre los valores pronosticados y cada variable explicativa no cambia cuando cambian las magnitudes de la variable explicativa, no hay heterocedasticidad en el modelo. Al calcular la prueba en ArcGIS, si se observa un asterisco próximo al Koenker P-Value, las relaciones que se están modelando presentan no-estacionariedad estadísticamente significativa (Rosenshein et al., 2011).

En la búsqueda para lograr una combinación de variables explicativas que produzcan coeficientes estadísticamente significativos, se deben ajustar los coeficientes al añadir y eliminar algunas de éstas. Los coeficientes pueden variar dramáticamente dependiendo de la combinación de variables que se incluyen en el modelo. Usualmente las variables que no son estadísticamente significativas son eliminadas del modelo.

Una vez se analiza el nivel de influencia de las variables explicativas al modelo en general, se debe analizar si las relaciones que el modelo muestra entre las variables explicativas y la variable dependiente son las esperadas (Rosenshein et al, 2011). No solo es importante determinar si una variable explicativa ayuda en la definición del modelo, también es importante reflexionar sobre la relación que tiene esta con la variable dependiente.

Cuando se seleccionan las variables explicativas, se espera que éstas respondan de manera positiva o negativa. Este patrón debe responder a patrones observados y a la lógica. Una relación negativa se puede dar cuando al modelar una condición se observa una variable explicativa cuyo valor del coeficiente asociado es negativo. Esto ocurre cuando el valor de la variable dependiente se reduce en la medida que la variable explicativa aumente (Burt et al., 2009). Un buen ejemplo de esta situación sucede con la edad del inmueble. A mayor cantidad de años, menor valor del coeficiente. En el caso de una relación positiva, en la medida que la variable independiente aumenta, igual aumenta la variable dependiente. Cuando se obtiene un resultado que no corresponde al planteamiento inicial, el modelo debe ser revisado y se deben aplicar otras pruebas estadísticas para corroborar el resultado.

Al momento de seleccionar las variables explicativas es importante evitar la elección de variables redundantes. Las variables explicativas deben ser independientes entre sí. Cuando se da esta condición se entiende que el modelo está ausente de multicolinealidad. Para conocer si se sufre de multicolinealidad se puede computar el Factor de Inflación de la

Varianza (FIV) para cada variable explicativa (Dziauddin et al., 2015). El resultado ayuda a decidir qué variables pueden ser eliminadas sin afectar el modelo.

Como regla práctica, las variables explicativas asociadas con los valores del FIV mayores que 7,5 aproximadamente deben eliminarse (una a una) del modelo de regresión hasta que haya logrado eliminar la redundancia (Zhou et al., 2015; Rosenshein et al., 2011). Antes de realizar la eliminación de la variable, se debe razonar sobre la eliminación de ésta. Por ejemplo, al modelar el valor de una residencia se puede contar con las variables explicativas de superficie del inmueble y cantidad de baños. Éstas podrían ser redundantes ya que en la medida que existan más baños se entiende que el inmueble ocupa más superficie. En este caso ambas pueden mostrar valores FIV altos. Se debe reflexionar sobre cuál de las variables eliminar del modelo.

Otro aspecto a observar en el proceso de diagnóstico es cuan bien ha sido especificado el modelo (Burt et al., 2009). Cuando se cuenta con un modelo MCO propiamente especificado, los residuales del modelo muestran una distribución normal con una media de cero. Cuando no hay parcialidad en el modelo, la distribución de los residuales no es equilibrada. Para determinar si se da esta condición o no, se puede utilizar el índice estadístico de Jarque-Bera (Smith et al., 2015). El índice estadístico de Jarque-Bera indica si los residuales (los valores de la variable dependiente observada o conocida, menos los valores previstos o estimados) se distribuyen normalmente o no. La hipótesis nula para esta prueba es que los residuales son distribuidos normalmente, por lo tanto, si se construyera un histograma de dichos residuales, su curva adoptaría una forma de campana. Cuando el valor p (probabilidad) de esta prueba es bajo, menor que 0,05 para un nivel de confianza del 95 por ciento, se entiende que los residuales no están distribuidos normalmente, lo que indica que el modelo es tendencioso. Si también cuenta con una autocorrelación espacial estadísticamente significativa de los residuales, la influencia puede ser el resultado de una especificación incorrecta del modelo, puede estar faltando una variable clave en el modelo (Burt et al., 2009). Los resultados de un modelo de MCO con especificación incorrecta no son confiables. También puede producirse una prueba de Jarque-Bera estadísticamente significativa si intenta modelar relaciones no lineales, si los datos incluyen valores atípicos influyentes o cuando hay heterocedasticidad marcada. Para analizar esto se pueden utilizar diagramas de dispersión. Con éstos se puede evaluar todas las relaciones entre las variables.

Además de la multicolinealidad y la heterocedasticidad, otro problema asociado a la aplicación de valoración hedónica a estudios de valores inmuebles es el de la autocorrelación espacial (Burt et al., 2009). Este es uno de los problemas menos atendido en la literatura (Suriatini, 2006). Los residuales producidos por un modelo de valoración hedónica aplicado a inmuebles pueden estar espacialmente correlacionados por tres razones (Suriatini, 2006; Tu et al., 2004; Gillen et al., 2001). Primero por las características estructurales del inmueble y las características del vecindario donde se localiza, segundo por el proceso de valoración utilizado para determinar el precio del inmueble y tercero por malas especificaciones del modelo MCO.

Para evaluar la autocorrelación espacial residual se puede calcular el Índice de Moran (Rosenshein, et al., 2011). Ésta es una estadística deductiva, lo que significa que los resultados del análisis siempre se interpretan dentro del contexto de la hipótesis nula. Para el Índice de Moran global, la hipótesis nula establece que el atributo que se analiza está distribuido en forma aleatoria entre las entidades del área de estudio; es decir, los procesos espaciales que promueven el patrón de valores observado constituyen una opción aleatoria. Por otro lado, el resultado de un valor de Z estadísticamente significativo indica que faltan variables explicativas muy importantes. No se puede confiar en los resultados de MCO cuando hay una especificación incorrecta del modelo. En el caso del ArcGIS, como resultado de la aplicación de los MCO se produce un mapa de residuales. Analizar este mapa puede ayudar a conseguir esas variables explicativas. Por defecto en éste, las zonas rojas indican que los valores observados son mayores a los pronosticados por el modelo. Las zonas azules indican que los valores observados son menores a los pronosticados.

Identificar qué valores ocurren por encima y por debajo de los observados puede ayudar a definir qué variables pueden estar faltando. Rosenshein et al. (2011) dan el ejemplo de sobre-predicción en las zonas urbanas. Si se observa esta tendencia, se puede introducir alguna variable que represente la distancia al centro urbano o alguna variable que represente la densidad urbana. Por otro lado, si las sobre-predicciones están asociadas a zonas montañosas o el fondo de algún valle, se debe considerar incluir alguna variable relacionada a la elevación del terreno. También se puede insertar una variable ficticia para capturar alguna diferencia regional. Por ejemplo, una variable que defina la zona urbana y la rural.

Otra prueba importante es el Criterio de información de Akaike corregido (AICc) (Burt et al., 2009; Rosenshein et al., 2011). Al comparar varios modelos aquel que presente el valor del

AICc más bajo es el mejor. En otras palabras, el modelo con el AICc más bajo se ajusta mejor a los datos observados.

Otro producto de la aplicación de los MCO en ArcGIS son las tablas de diagnóstico y coeficientes para los modelos. Estos capturan los elementos más importantes del informe de MCO, incluida la lista de variables explicativas utilizada en el modelo con los coeficientes, los errores estándar y las probabilidades, así como los resultados para cada prueba de diagnóstico. La tabla de diagnóstico incluye una descripción de cada prueba junto con algunas pautas para interpretar los resultados de las pruebas.

Finalmente se debe evaluar el rendimiento del modelo (Suriatini, 2006; Smith, et al., 2007; Liu et al., 2010; Rosenshein et al., 2011). Los valores del Coeficiente de determinación múltiple R^2 y R^2 ajustado son medidas del rendimiento del modelo. Los valores posibles varían de 0,0 a 1,0. El valor R^2 ajustado siempre es un poco más bajo que el valor del Coeficiente de determinación múltiple R^2 , porque refleja la complejidad del modelo (la cantidad de variables) ya que se relaciona con los datos y es, por lo tanto, una medida más exacta del rendimiento del modelo. Si se agrega una variable explicativa adicional al modelo, probablemente el valor del Coeficiente de determinación múltiple R^2 aumente y el valor R^2 ajustado disminuya. Un valor R^2 ajustado de 0,84 indicará que el modelo, sus variables explicativas modeladas con una regresión lineal, explica aproximadamente el 84 por ciento de la variación en la variable dependiente. En otras palabras, el modelo proporciona aproximadamente el 84% de la información sobre robo residencial.

3.3 Regresión Geográficamente Ponderada

Al revisar la literatura se puede establecer que las primeras aplicaciones del modelo de valoración hedónica para analizar precios de inmuebles utilizaban como fundamento los métodos y estadísticas de la econometría tradicional (Ridker y Henning, 1967). Como se ha discutido antes, ésta excluye el espacio como fuente espacial de los datos al no tener en cuenta la correlación espacial de estos (Anselin, 1998). Esto trae consigo dos problemas, el de la correlación espacial y el de la heterogeneidad espacial. La correlación espacial se presenta cuando hay niveles de dependencia espacial entre las variables. Esto se evidencia en los mercados inmobiliarios cuando se cuenta con residencias de diversos valores dentro de un mismo espacio geográfico. Por otro lado, el problema de la heterogeneidad espacial presenta

una inestabilidad espacial de los coeficientes obtenidos por el modelo. Esto se ejemplifica cuando se analizan diversos sectores inmobiliarios y se observa que una característica es importante en un sector, no siéndolo en otro.

Para superar esa condición, Brunsdon et al. (1996) introducen el método de la “Regresión Geográficamente Ponderada” (*Geographic Weighted Regression*). Para ellos la no estacionariedad espacial es una condición en la que un modelo simple global no puede explicar las relaciones existentes entre algunos conjuntos de variables. Por tal razón la naturaleza del modelo a utilizar debe alterar el espacio donde se localizan las variables para poder reflejar la estructura interna de los datos. Con la Regresión Geográficamente Ponderada, Brunsdon et al. (1996) intentan captar esta variación, calibrando un modelo de regresión múltiple que permite la existencia de diferentes relaciones en diversos puntos en el espacio.

La Regresión Geográficamente Ponderada permite coeficientes de regresión que varían a través del espacio fundamentados en el postulado de Tobler (1970). En el marco de la Regresión Geográficamente Ponderada, la fórmula del modelo hedónico tradicional puede ser reescrita como:

$$P_i(H) = f_i(S, N) + \varepsilon_i$$

donde el subíndice i representa lugares geográficos específicos y los otros símbolos se definen como en la ecuación hedónica original. Esta ecuación indica que potencialmente, en diferentes posiciones geográficas, la ecuación hedónica puede tomar diferentes formas (Yu, 2007). En una especificación lineal, la ecuación se puede simplificar como:

$$P_i(H) = \beta_{i0} + \beta_{iS}S_i + \beta_{iN}N + \varepsilon_i$$

donde β_{i0} es el término de intersección, β_{iS} y β_{iN} son coeficientes de características estructurales y de vecindad que varían espacialmente.

La calibración del modelo de Regresión Geográficamente Ponderada sigue un enfoque de mínimos cuadrados ponderados localmente (Fotheringham et al., 2002). A diferencia de los mínimos cuadrados ordinarios, al calibrar los coeficientes en la localización i , la Regresión Geográficamente Ponderada asigna pesos a través de un esquema de ponderación a los datos en las ubicaciones de acuerdo a su proximidad espacial a la posición i (Yu, 2007). Estos pesos actúan para asegurar que los lugares más cercanos ejercen una mayor influencia que

aquellos lugares más alejados. Los pesos se obtienen generalmente a través de una función del núcleo (Kernel) espacial. Dos tipos de núcleos espaciales se utilizan, los fijos y los adaptativos (Fotheringham et al., 2002). Con el núcleo fijo se obtiene un kernel espacial óptimo el cual se aplica sobre el área de estudio. Este enfoque suele ser menos intenso computacionalmente. Sin embargo, como ha señalado Fotheringham et al. (2002) y Tu et al., (2004), el enfoque del núcleo fijo puede producir gran varianza de estimación local en áreas donde los datos son escasos y puede enmascarar variaciones locales sutiles en áreas donde los datos son densos. Por otro lado, el núcleo de adaptación busca un cierto número de vecinos más cercanos para adaptar el kernel espacial y asegurar un tamaño constante de las muestras locales. Este núcleo puede presentar medios más razonables para representar el grado de no estacionariedad espacial en el área de estudio.

Para obtener el tamaño óptimo de los vecinos más cercanos a ser utilizado con el núcleo de adaptación, un enfoque común es minimizar el Criterio de Información de Akaike (AIC) del modelo de Regresión Geográficamente Ponderada (Fotheringham et al., 2002). El AIC de un modelo de Regresión Geográficamente Ponderada se define siguiendo las obras de Hurvich (Tu et al., 2004):

$$AIC = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n \left(\frac{n + \text{tr}(S)}{n - 2 - \text{tr}(S)} \right)$$

donde n es el número total de observaciones, $\hat{\sigma}$ es la probabilidad máxima estimada de la desviación estándar del término de error, y $\text{tr}(S)$ es la traza de la matriz S de la Regresión Geográficamente Ponderada, que se define como:

$$\hat{y} = Sy$$

donde y y \hat{y} son el vector de la variable dependiente y los valores estimados de la Regresión Geográficamente Ponderada. El AIC tiene el atractivo general, que podría ser utilizado para evaluar si la Regresión Geográficamente Ponderada proporciona un mejor ajuste de un modelo global teniendo en cuenta los diferentes grados de libertad en los dos modelos.

Aunque la técnica de Regresión Geográficamente Ponderada parece ser atractiva para analizar la no estacionariedad espacial, la estrategia de modelado y procedimiento de calibración en estas son más complejos que en un modelo global de Mínimos Cuadrados Ordinarios.

Desde su introducción, este modelo ha sido aplicado a estudios de diversas índoles. Cuando Brunsdon et al. (1996) presentan el modelo, lo ponen a prueba con un conjunto de datos donde se relaciona a dueños de automóviles con su clase social y la tasa de desempleo. Los datos utilizados se obtienen del censo del Reino Unido para el año 1991. En el caso del estudio de valores residenciales son varias las aplicaciones que se han realizado de la Regresión Geográficamente Ponderada. Por ejemplo, Brunsdon et al. (1999) utilizan el modelo para describir la distribución espacial de los valores residenciales en la ciudad de Kent, en el Reino Unido.

Para examinar el método, Bitter et al. (2007) buscaron determinar la heterogeneidad espacial de valores residenciales al comparar el método de la Regresión Geográficamente Ponderada contra el Método de Expansión Espacial con datos de la ciudad de Tucson en Arizona. Con su estudio los autores demuestran claramente la distribución espacial de variables claves en la predicción de los valores residenciales siendo la Regresión Geográficamente Ponderada la cual muestra mejores resultados explicativos y mejor precisión en la predicción. Por su parte Yu (2007) estudió la relación espacial no estacionaria entre los valores residenciales y sus respectivos atributos en la ciudad de Milwaukee. La Regresión Geográficamente Ponderada reveló la existencia marcada entre el valor de la residencia y las variables predictivas. Encontró en particular que la superficie de ocupación, el número de cuartos de baño y la existencia de acondicionadores de aire y chimeneas presentan una influencia mayor en el valor. En una aplicación más reciente, Lu et al. (2014) analizan la relación de valores residenciales con diferentes características en la ciudad de Londres. Ellos parten del principio que cada regresión local en la Regresión Geográficamente Ponderada se estima con datos cuya influencia decae con la distancia y las distancias se establecen como una línea recta o euclidiana. Los autores comparan tres aplicaciones de la Regresión Geográficamente Ponderada cuya diferencia estriba en la forma en que calibran los modelos. Los modelos se calibran con la distancia euclidiana, distancia de la red de carreteras y métricas del tiempo de viaje. Los resultados indican que al utilizar distancias no euclidianas, los resultados de la Regresión Geográficamente Ponderada mejora el ajuste del modelo.

Capítulo 4

Caguas: Una Mirada a su Desarrollo Urbano

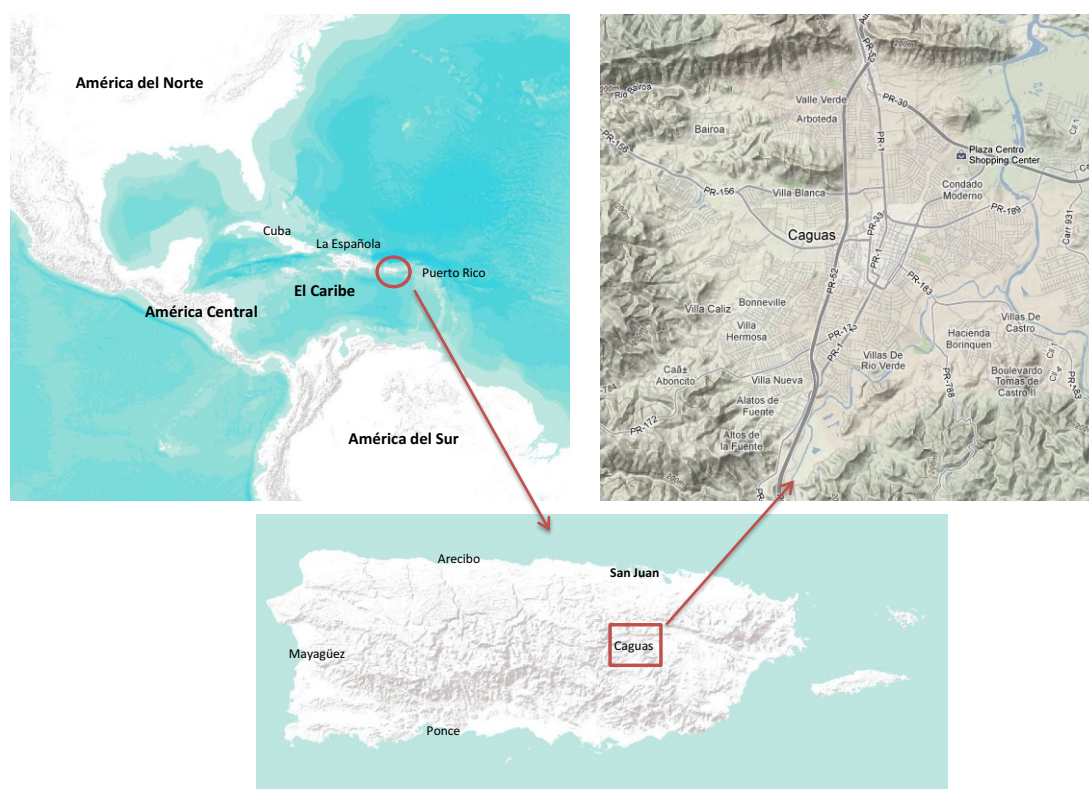
4. Caguas: Una Mirada a su Desarrollo Urbano

4.1 Localización

Para este estudio se seleccionó el municipio de Caguas, Puerto Rico (Figura 4.1). El municipio de Caguas está localizado en la porción central este de Puerto Rico, contando con una extensión aproximada de 152 km² (un 1,7% del territorio de Puerto Rico). Caguas se localiza al sur de San Juan (Capital de Puerto Rico) y su Región Metropolitana.

Sus características fisiográficas principales son: las tierras planas y semiplanos del valle central, las colinas al norte del valle y la zona más escarpada hacia el sur y el oeste. El valle ocupa aproximadamente 41 km² y se encuentra a una elevación promedio de 60 m.s.n.m.. Las colinas al norte del valle cuentan con elevaciones entre los 90 y 240 m. y las zonas más escarpadas al sur y oeste con elevaciones máximas de 600 m.s.n.m.

Figura 4.1: Localización de Puerto Rico y Caguas.



Fuente: Elaboración Propia

4.2 Evolución Urbana

4.2.1 De los orígenes a finales del siglo XX

Una simple mirada a las fotografías aéreas existentes en Puerto Rico desde 1940 permite afirmar que el proceso de urbanización en Puerto Rico ha ocurrido de manera acelerada. La transformación de una sociedad que hace apenas unos 70 años se basaba en la agricultura, cuya población se asentaba en las zonas rurales, a una sociedad predominantemente urbana, ocurrió en un espacio de tiempo muy corto. Se puede entender que debido a lo acelerado de este proceso propició la morfología urbana que impera actualmente. A esto se debe sumar que el modelo de urbanización adoptado es el de los Estados Unidos, un territorio continental con vastos espacios donde el establecimiento de vivienda unifamiliar no solo se amoldaba a este entorno, si no que se ajustaba perfectamente al modelo de producción industrial de la post-guerra. Sin embargo, este modelo no se ajusta a la realidad de Puerto Rico.

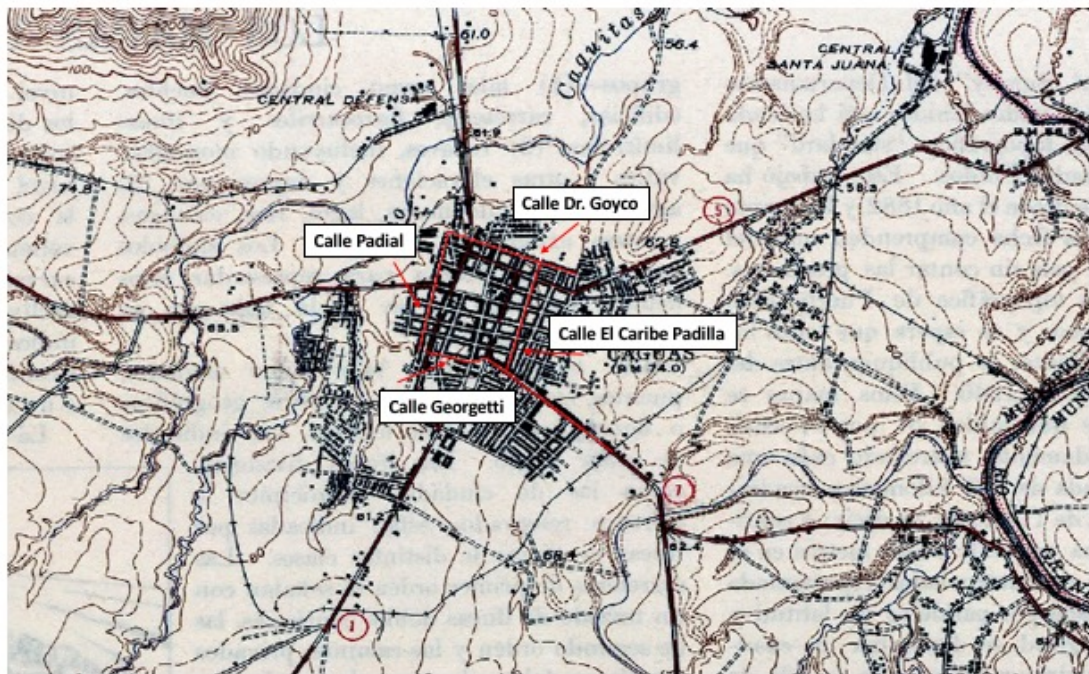
Las ciudades puertorriqueñas muestran dos patrones morfológicos urbanos muy marcados: el reticular de los centros tradicionales que corresponden al período de dominio español; y el segundo, en el que domina el modelo de vivienda unifamiliar, traslación del modelo norteamericano de construcción de vivienda suburbana, en el que se da muy poco orden y no existe relación lógica entre los usos.

Para ver estos patrones en el caso de Caguas es fundamental estudiar su evolución urbana. La región de Caguas fue un gran asentamiento indígena pre-colombino. Aunque el origen del patrón urbano actual data del año 1775, no se puede obviar de la historia el hecho que los indios Tainos ya contaban con un gran asentamiento en la zona (Bunker, 1983). En su faena de exploración, los colonizadores se toparon con el asentamiento del Cacique Caguax. La riqueza de los suelos y cuerpos de agua hizo de este un espacio geográfico de alto valor para establecer un asentamiento para los indígenas. El nombre de Caguas se deriva del nombre de este cacique.

El asentamiento que se conoce hoy data de 1775, en el cual residían 722 habitantes (Plan de Ordenación Territorial Caguas, 2011). En el año 1814 se comienza una reorganización de este espacio en torno a la plaza central del pueblo. Con el pasar de los años se fue estableciendo en torno a la plaza las diversas calles y usos residenciales y de comercio. Para la época la economía predominante era la agrícola, por lo que gran parte de los comercios que se establecían giraban en torno a esta industria. En 1894 Caguas obtiene el título de ciudad (Bunker, 1983).

En 1898, para la época de la llegada de las tropas militares de Estados Unidos de Norteamérica, Caguas contaba con 17.820 habitantes (Plan de Ordenación Territorial Caguas, 2011). La zona urbana se delimitaba por las calles Dr. Goyco (norte), Georgetti (Sur), Padilla (El Caribe) (Este) y Padial (Oeste) (Ver Figura 4.2). Para la época las calles no estaban pavimentadas ni existían servicios básicos. Los edificios alrededor de la plaza eran de mampostería; fuera de esa zona eran de madera techados en tejas de barro o zinc galvanizado (Ver Imagen 4.2). El resto del territorio del valle estaba compuesto por estancias para la siembra de café, caña de azúcar y tabaco (Ver Imagen 4.1). Entre estas se encontraba la Hacienda Santa Catalina, la Hacienda San José, La hacienda Cuatro Calles y la Hacienda Santa Bárbara (Acosta, 1995).

Figura 4.2: Caguas, 1898



Fuente: United States Geological Survey, 1946

Imagen 4.1: Siembra de Tabaco en las afueras de ciudad



Fuente: Librería del Congreso de los Estados Unidos (<http://www.loc.gov/pictures/>) (Accedido 19 de noviembre del 2015)

El siglo XX trae cambios dramáticos al modelo económico del pueblo con una influencia directa en su crecimiento urbano. Bunker (1975 y 1983) indica como fechas importantes para el desarrollo de infraestructura en el municipio el año 1908, cuando se establece el sistema de alumbrado eléctrico en torno la plaza central y se inaugura el acueducto del pueblo, y el año 1912, cuando se inaugura el sistema de alcantarillado.

Para el 1908, se observa el primer ensanche fuera del tejido original del pueblo. El ayuntamiento adquiere un predio de terreno al sureste del pueblo con el propósito de urbanizarlo. Se desarrolla allí la primera urbanización de Caguas llamada Campo Alegre. La misma consistía de 24 residencias. Más tarde, en 1912, al Sureste de la ciudad se desarrolla la primera urbanización privada en Caguas. Esta fue desarrollada por los señores Harrison Johnson y Pablo Héreter. La misma consistía de 300 solares en el sector conocido actualmente como Savarona (Plan de Ordenación Territorial de Caguas, 2011).

Imagen 4.2: Calle Principal. Se puede apreciar la misma sin pavimentar. Casas de madera techados en tejas de barro.



Fuente: Librería del Congreso de los Estados Unidos (<http://www.loc.gov/pictures/>) (Accedido 19 de noviembre del 2015)

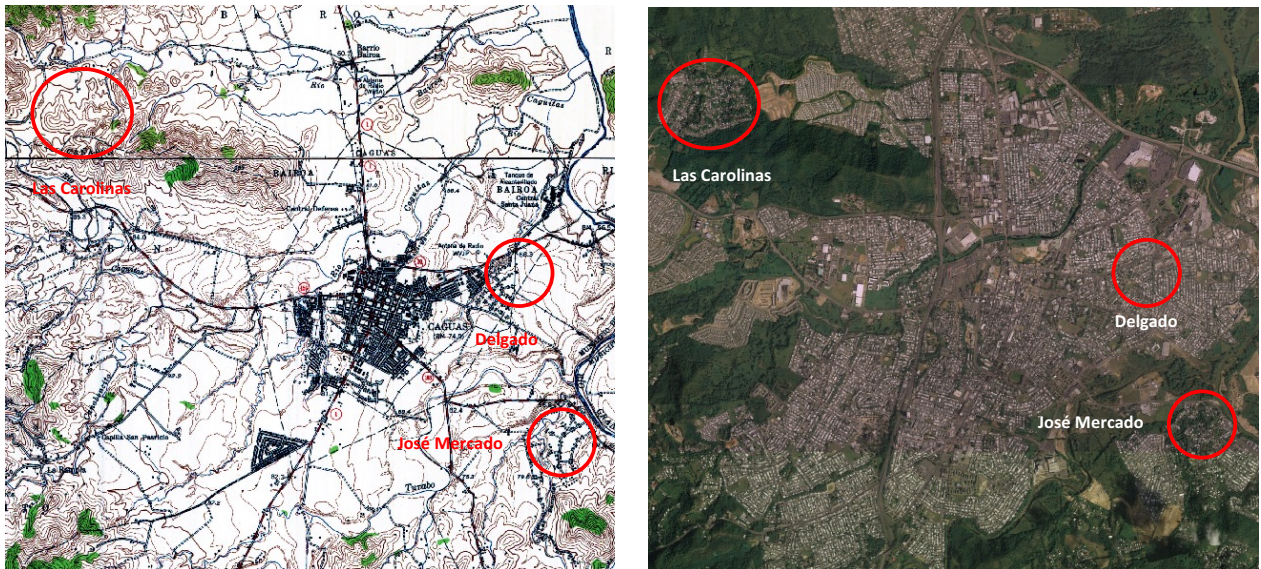
Todo el crecimiento se redujo dramáticamente con la Gran Depresión de la década del 1930. Esta dio un duro golpe al crecimiento económico de la ciudad. Franklin Delano Roosevelt se convierte en presidente de los Estados Unidos e impulsa un plan de renovación nacional centrado en la generación de obra pública para incentivar el empleo. Este plan, mejor conocido como “El Nuevo Trato” dio paso en Puerto Rico a la creación de la “Puerto Rico Emergency Relief Administration”, o PRERA por sus siglas en inglés. Uno de los productos de este plan fue la creación de la primera planta de cemento en Puerto Rico. Con esto se buscaba incentivar la industria de la construcción. Caguas vio el beneficio de esto con la pavimentación de las calles y creación de encintados y aceras en el centro del pueblo.

Además, se mejoraron las carreteras que conectaban a Caguas con la capital (San Juan) y los pueblos limítrofes.

La PRERA también generó un programa de eliminación de asentamientos informales. Surge en Puerto Rico el concepto de vivienda para personas de escasos recursos. En Caguas el primer ejemplo es la construcción de la urbanización La Granja (1936). Para la década del 1940 surge la construcción de proyectos de vivienda rural. Estos proyectos se desarrollaban en las zonas rurales aledañas al pueblo. Dado el crecimiento de la ciudad hoy día muchos de estos quedan entremezclados dentro del tejido de la ciudad. Como ejemplo se pueden mencionar la comunidad Las Carolinas, Muñoz Grillo, La Jurado, La Barra, José Mercado y José Delgado (Figura 4.3).

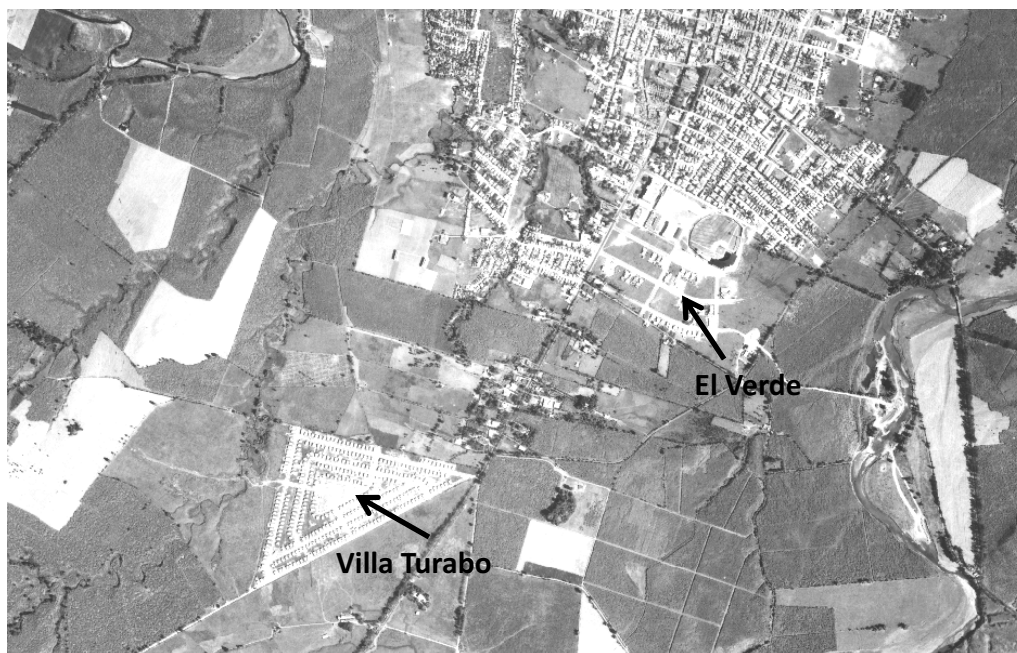
El concepto moderno de urbanización llega a Puerto Rico con la construcción de la urbanización Puerto Nuevo en San Juan. Este desarrollo comenzó su construcción en 1947. Este es el primer desarrollo de vivienda unifamiliar alejado del patrón reticular heredado del período español. El nuevo modelo se refleja en Caguas en el 1950 con la construcción de la urbanización Villa Turabo (Imagen 4.3). Este se construyó a una distancia considerable de lo que era el centro del pueblo. Este se convirtió en el modelo a seguir por los desarrolladores en Puerto Rico. Este fue el primer caso de sub-urbanización en Caguas. En 1951 se construye la urbanización El Verde (Imagen 4.3). Esta se ubica al sur de la zona urbana y constituye la primera urbanización dirigida a familias de alto poder adquisitivo. En esta década se construye la urbanización Villa Blanca y un parque industrial al oeste de esta urbanización. Ambos desarrollos se localizan al norte del centro urbano. De esta manera se siguen desarrollando urbanizaciones de manera dispersa sin seguir un patrón definido.

Figura 4.3: Comunidades rurales creadas bajo el programa de la PRERA. A la izquierda localizadas sobre mapa topográfico de 1950 y a la derecha su localización actual.



Fuente: Mapa Topográfico Servicio Geológico de los Estados Unidos, 1950 y Fotografía aérea producida por la Autoridad de Carreteras de Puerto Rico.

Imagen 4.3: Urbanizaciones Villa Turabo y El Verde (1951)



Fuente: Fotografía aérea producida por la Autoridad de Carreteras de Puerto Rico, 1951

Imagen 4.4: Casa de Villa Turabo Actualmente



Fuente: Raúl Matos Flores, 2015

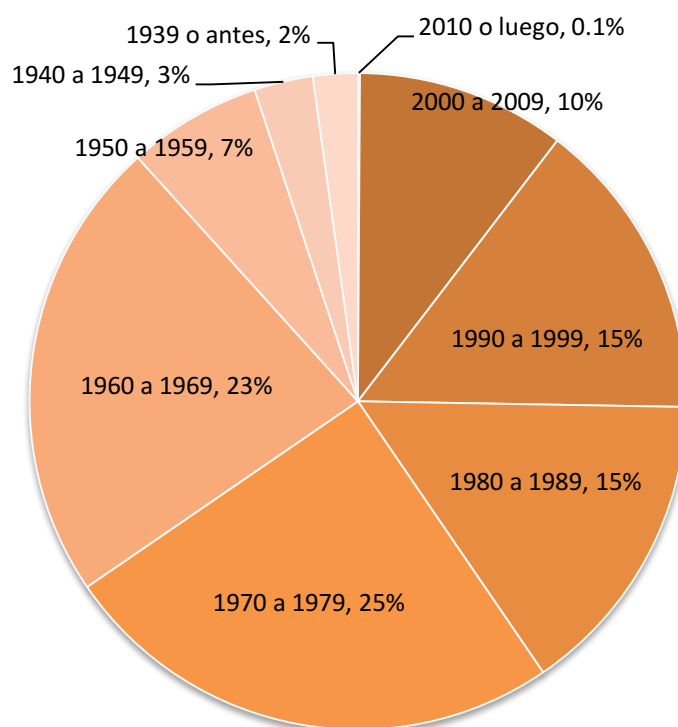
Con la década del 1960 comienza la transformación de una economía basada en la agricultura a una basada en el nuevo modelo industrial (Bunker, 1983). Desaparece el uso agrícola que se daba a los terrenos circundantes a la ciudad y estos poco a poco se van transformando en terrenos dedicados a la construcción de viviendas. Entre 1961 y 1968 se construyeron en Caguas 12.978 residencias. Entre 1960 y 1972 se añadieron 45 km² (2,173 cuerdas) de espacios construidos. Esta cifra representa tres veces la superficie ocupada para usos urbanos en el 1960.

La década del 1970 sufre una reducción en el proceso de urbanización debido principalmente a la desaceleración de la economía que duró desde principios de la década del 1970 hasta principios de la de 1980. Los desarrollos que se realizaron fueron todos de tipo unifamiliar. La actividad de construcción de viviendas acelera nuevamente a mediados de la década del 1980 estableciendo urbanizaciones más alejadas del centro de la ciudad y cuyo perfil de venta va dirigido a familias de poder adquisitivo medio a alto.

En la década del 1990 el proceso de urbanización se dedica a urbanizar espacios sin urbanizar dentro del tejido urbano. Además de la típica residencia unifamiliar se observa la proliferación de desarrollos multifamiliares llamados “walk-ups”⁵.

Gráfica 4.1

Década de Construcción de Viviendas



Fuentes: Negociado del Censo federal, Censo Decenal de Población.

4.2.2 Caguas Contemporáneo

Como se pudo apreciar en la sección pasada, Caguas se ha caracterizado en la segunda mitad del siglo pasado por dos marcadas tendencias poblacionales: el aumento constante de su población y su transformación en municipio predominantemente urbano. Este aumento en población responde en parte a la mejora en la economía de la ciudad, pero también se debe al proceso migratorio de la Región Metropolitana de San Juan a Caguas. Los altos costos de

⁵ Edificio de Apartamentos que no cuenta con ascensor. No supera las 4 plantas.

adquisición de vivienda en la Región Metropolitana de San Juan provocaron que muchas familias se trasladaran hacia la periferia en busca de vivienda más accesible. Ante una oferta de vivienda a costos moderados en Caguas durante las décadas de 1970 y 1980, muchas familias que laboraban en la Región Metropolitana de San Juan decidieron establecer sus residencias en Caguas. A esto se suma la construcción de la autovía que conecta San Juan con Caguas a finales de la década del 1970.

Tabla 4.1 – Población por década

Año	1950	1960	1970	1980	1990	2000	2010
Población	60.132	65.098	95.661	117.959	133.447	140.502	142.893

Fuente: Negociado del Censo federal, Censo Decenal de Población

Este desplazamiento de la población combinado con el crecimiento natural de la población, resultó en un aumento en la demanda por vivienda. Este patrón se sostuvo hasta finales de la década del 1990. Ya en la primera década del nuevo siglo son los municipios periféricos a Caguas los que comienzan a ofrecer nuevos desarrollos residenciales con altos atractivos en términos de facilidades recreativas y seguridad a costos menores de los que se pueden encontrar en Caguas. Al igual que muchas personas vieron a Caguas como atractivo entre las décadas de 1970 al 1990, ahora son los municipios periféricos a Caguas los que ejercen dicha atracción. Esto se puede ver reflejado en las cifras de población para estos períodos (Tabla 4.1). La población de Caguas apenas aumento 2.391 personas del 2000 al 2010, un crecimiento de un 1,7 por ciento.

Tabla 4.2 – Porcentaje de cambio de población década a década.

1950 / 1960	1960 / 1970	1970 / 1980	1980 / 1990	1990 / 2000	2000 / 2010
8,3	46,9	23,3	13,1	5,3	1,7

Fuente: Negociado del Censo federal, Censo Decenal de Población

Según el Censo de Población y Vivienda de 2010, Caguas tiene 60.371 unidades de vivienda (Tabla 4.3). Esto representa un aumento de 13.434 viviendas en la última década. De las viviendas contabilizadas en el 2010, un total de 7.098 se encontraban desocupadas (Tabla 4.3). Al comparar con las pasadas décadas esta cifra muestra la mayor relación de viviendas ocupadas / desocupadas de los pasados 60 años. El promedio de valor de la vivienda en Caguas fue de \$142.600. Esto representa una diferencia en aumento de \$56.900 con respecto al promedio del año 2000 (Tabla 4.4).

Tabla 4.3 – Vivienda por décadas

Año	1950	1960	1970	1980	1990	2000	2010
Viviendas	12.301	14.420	24.904	35.226	43.293	50.568	60.371
Ocupadas	11.638	13.264	22.291	31.903	40.104	46.937	53.273
Desocupadas	663	1.156	2.613	3.323	3.189	3.631	7.098

Fuente: Negociado del Censo federal, Censo Decenal de Población

Tabla 4.4 – Valor de Viviendas Ocupadas (En dólares)

Década	Viviendas	Valor Promedio (Dólares)
1970	22.291	12.000
1980	31.903	26.200
1990	40.104	44.600
2000	46.937	85.700
2010	53.272	142.600

Fuentes: Negociado del Censo federal, Censo Decenal de Población.

La última década se ha caracterizado por la proliferación de proyectos nuevos de vivienda unifamiliares de alto costo. Se construyen viviendas para sectores de alto poder adquisitivo cuyo costo es elevado, pero no comparado con el de la Región Metropolitana de San Juan. Los valores residenciales alcanzaron cifras nunca registrados en la historia. Como contraparte, se estima que a partir del año 2007 Puerto Rico se encuentra inmerso en una recesión económica. Este es un periodo durante el cual el desempleo ha aumentado

significativamente y el ingreso del país se ha reducido un 10%. Uno de los sectores afectados por este ambiente económico es el mercado inmobiliario. Es normal observar una reducción en la demanda por vivienda y que junto a esto los bancos estén menos dispuestos a financiar la compraventa de vivienda. Ambos factores generan presión hacia una reducción en el mercado inmobiliario. Además, se registra una reducción significativa en el valor de las viviendas. Marxuach (2010) establece que esa reducción se estima en un 25%, lo que representa una pérdida agregada estimada de alrededor de 27.000 millones de dólares.

Sin embargo, Marxuach (2010) plantea que la magnitud de las pérdidas reportadas no se puede considerar solo como un efecto de la recesión. Entiende Marxuach que existe un serio problema estructural en el mercado de vivienda en Puerto Rico. Para entender esto se deben analizar los datos censales más recientes. La información recopilada por el “Puerto Rico Community Survey” de la Oficina del Censo muestra una reducción en el número de hogares⁶ en Puerto Rico. Entre el 2005 y el 2009 la cantidad total de hogares en Puerto Rico se redujo de 1.254.318 a 1.181.112. Esta es una reducción de 73.206 hogares, equivalente a un 5,8%. Esta merma en la cantidad de hogares se relaciona con una reducción en la demanda de vivienda. Para este período la Junta de Planificación de Puerto Rico reportaba aumentos en la construcción de nuevas residencias. Las nuevas residencias se construían y valoraban con la libertad que el mercado hipotecario y los bancos habían creado. Además, muchas de esas unidades residenciales eran para el mercado de vivienda secundaria. No lograr que estas residencias fueran vendidas creó un exceso de residencias en un mercado de poca demanda. El desfase entre la construcción y la oferta de viviendas era evidente.

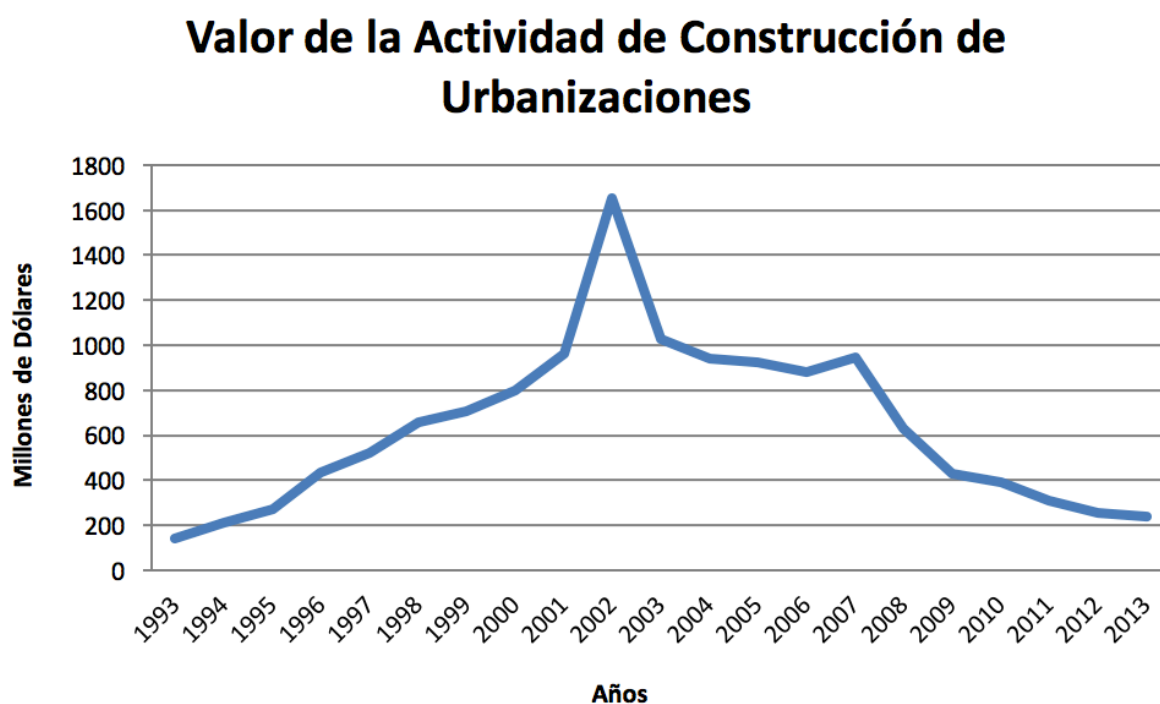
Este desfase evidenció la existencia de una burbuja en el sector de bienes raíces cuyo punto crítico ocurrió entre el año 2006 y el año 2008. No había una demanda real para todas esas viviendas, pero los precios seguían subiendo fuera de toda proporción con los niveles de ingreso del país. Era muy difícil que los precios de los inmuebles bajaran. Esta es una de las señales clásicas de una burbuja: los precios subían simplemente porque todo el mundo esperaba que subieran.

La burbuja fue facilitada por una oleada de préstamos de construcción y desarrollo de tierras. De acuerdo con el informe “Estadísticas Seleccionadas sobre la Industria de la Construcción

⁶ La oficina del Censo distingue entre vivienda y hogar. Se contabiliza como hogar cuando la vivienda fue adquirida por un individuo o familia y la misma es habitada. De lo contrario se considera una vivienda que está desocupada. Entre las razones pueden estar el que la vivienda no ha sido vendida por el desarrollador o que la familia abandono el inmueble por alguna razón (proceso de embargo de parte de un banco o emigración).

“para los años 2000, 2007 y 2013, el pico de la inversión económica para la construcción de nuevas urbanizaciones ocurrió entre los años 2000 al 2003 (Gráfica 4.2). Para el año 2002 ésta generó un total de 1.654 millones de dólares. Luego del estallido de la burbuja inmobiliaria situado entre 2007 y 2008 se puede apreciar como esta cifra baja dramáticamente. En el año 2007 la inversión se encontraba en 945 millones de dólares; esta cifra se reduce a 627 millones de dólares en el año 2008 y a 427 millones de dólares en el año 2009. Esto representa una reducción de 518 millones de dólares en dos años. Al año 2013 la cifra se estimó en 237 millones de dólares. Este punto climático de la crisis financiera mundial tiene su impacto en el sistema financiero local. El exceso de residencias sin vender se mantiene a la fecha. Muchas de éstas se consideran mansiones. Esto solamente por su costo de venta ya que en términos de edificación son similares a residencias construidas en la década pasada. Se concluye que el costo asignado por el desarrollador solo era efecto de la burbuja inmobiliaria.

Gráfica 4.2 Valor Actividad de Construcción



Fuente: Estadísticas Seleccionadas sobre la Industria de la Construcción, Junta de Planificación de Puerto Rico

Imagen 4.5: Ejemplo nuevas casas para sector económico de alto poder adquisitivo



Fuente: <http://axesabienesraices.com> (Accedido 14 enero 2011)

4.3 Tipología edificatoria en Caguas

Los sub-mercados residenciales se definen como áreas geográficas donde el precio por unidad de vivienda es constante debido a la existencia de características comunes de las viviendas. Para la segmentación del mercado residencial se puede utilizar como criterios las diferencias espaciales de las características estructurales de los inmuebles, los servicios con que cuenta el vecindario, o una combinación de ambos. Por ejemplo, residencias construidas en años diferentes pueden tener características únicas que no son fácilmente cambiables o reproducibles. Para muchos demandantes de estas residencias su idea particular de un hogar no puede ser sustituible y los costos de una posible transformación pueden ser sustanciales. Además, las amenidades y facilidades de un vecindario no son fáciles de duplicar.

Imagen 4.6: Caguas en la actualidad



Fuente: <http://www.geocities.ws/caguasmipueblo/FotoCaguasAerea.htm> (Accedido 20 de septiembre 2014)

En el caso de Caguas y la tipología edificatoria unifamiliar domina ampliamente el paisaje urbano limitando a pocos casos las estructuras de viviendas plurifamiliares (Mapa 4.5).

Vivienda unifamiliar es aquella en la que una única familia, persona o grupo de personas, viven o no, en determinadas condiciones, en una construcción normalmente de una a dos plantas y con espacio natural (zona verde) circundando la vivienda. Estas se dividen en tres modalidades:

- Unifamiliar aislada: Edificio habitado por una única familia que no está en contacto físico con otras edificaciones. Normalmente están rodeadas por todos sus lados por un terreno perteneciente a la vivienda, en el que se suele instalar un jardín privado. En este aspecto hay variantes; así, la vivienda puede tener uno, varios o todos sus lados alineados con la vía pública.

- Unifamiliar pareada: En este caso, se construyen dos viviendas unifamiliares que exteriormente están en contacto, aunque en su distribución interior son totalmente independientes, teniendo cada una de ellas su propio acceso desde la vía pública. También se les conoce como casas Dúplex.
- Unifamiliar adosada: Similar a la pareada, pero esta vez cada vivienda está en contacto con otras dos (una a cada lado). Este tipo de viviendas se suelen caracterizar por tener una planta estrecha y alargada y por la presencia de ventanas únicamente en los extremos de la casa.

Imagen 4.7: Unifamiliar aislada



Fuente: <http://clasificadospr.com/v/bienes-raices-residencial-casas-multi-familiares/puerto-rico-caguas/excelente-residencia-mirador-de-bairoa-526250s> (Accedido 21 de septiembre 2014)

Imagen 4.8: Casa adosada de una planta



Fuente: Raúl Matos Flores (2016)

Imagen 4.9: Adosada de dos plantas.



Fuente: Raúl Matos Flores (2016)

Los complejos de vivienda multifamiliar se componen de uno o varios edificios de uso mayoritariamente residencial que dispone de acceso y servicios comunes para más de dos viviendas. También es conocido como vivienda plurifamiliar. En el caso de Caguas se pueden observar dos modalidades:

- Condominios: Se le conoce como condominio en Puerto Rico a las edificaciones verticales que superan las 5 plantas y que cuentan con ascensor como medio de acceso a las plantas superiores.
- Vivienda tipo “Walk-up”: Edificio de apartamentos de no más de 4 plantas en donde no existe ascensor

Imagen 4.10: Caguas Tower



Fuente: Raúl Matos Flores (2016)

Imagen 4.11: Condominio Santa Juana



Fuente: Raúl Matos Flores (2016)

Imagen 4.12: Desarrollo tipo “Walk-up”



Fuente: <http://www.compraalquila.com.pr/proyectos-nuevos/valle-santa-cecilia/puerto-rico/caguas/53/> (Accesado 15 diciembre 2016)

Independiente de la tipología edificatoria, a partir de finales de la década del 1980 las nuevas urbanizaciones comienzan a construirse con controles de acceso. En el 1987 se aprueba la “Ley de Control de Acceso Para Urbanizaciones de Puerto Rico”. El concepto de la urbanización cerrada comienza a dominar el paisaje urbano.

Una urbanización cerrada (del inglés “gated community”) es una forma especial de barrio residencial, cuyo ordenamiento y vialidad puede ser orden privado, con un perímetro definido por muros o rejas, y posee una única entrada controlada por un servicio de seguridad, los cuales se encargan de comprobar la identidad de los visitantes y anunciarlos.

Una urbanización cerrada tiene por lo general pequeñas calles residenciales, con sus propias normas de tránsito. Las urbanizaciones cerradas fueron desarrolladas para satisfacer el deseo de seguridad y prestigio de sus compradores. Los muros que las rodean, en muchos casos con cámaras de vigilancia, dan a sus habitantes un sentido de seguridad.

Un aspecto importante bajo el concepto de las urbanizaciones cerradas es el de acceso a las calles. En Puerto Rico las calles son bienes del dominio público. Es ilegal privar el uso de ellas a los habitantes que no residen en la urbanización. En otros países bajo el concepto de la urbanización cerrada las calles son propiedad de los residentes de la urbanización y con esto adquieren la responsabilidad de mantenimiento de estas. En Puerto Rico éste no es el caso. Los residentes asumen los costos de seguridad privada, mantenimiento de facilidades recreativas y paisajismo, mientras el estado ofrece servicios como recogido de basura y mantenimiento de calles.

El hecho de estar en una comunidad cerrada da cierto nivel de exclusividad. Esta condición hace que los desarrolladores al momento de diseñar el complejo residencial generen áreas de uso común o “*amenities*”, tales como casa club (clubhouses), áreas comerciales, canchas deportivas, piscinas, parques infantiles, campos de golf, gimnasios, escuelas, y hasta helipuertos. Esto permite establecer como objetivo de venta a personas o familias de clase media-alta hasta las de clases altas. Como consecuencia, el costo de las propiedades aumenta en la medida que las áreas de uso común que se ofrecen con la comunidad apelen al gusto de la clase económica al que se tiene como objetivo. Al ser una comunidad cerrada será el residente quien aporte las cuotas anuales o mensuales para el mantenimiento de las facilidades. Esta aportación también va en función del tipo de facilidad. La cuota para mantener un parque infantil no será igual al que cuenta con un campo de golf.

Imagen 4.13: Ejemplos de Control de Acceso

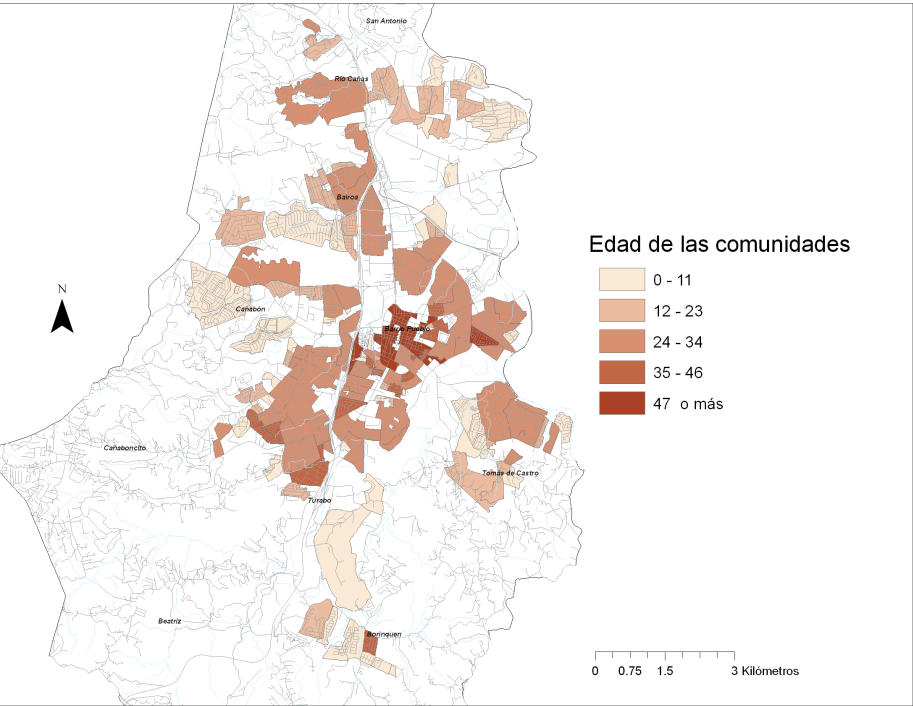


Fuente: Raúl Matos Flores (2016)



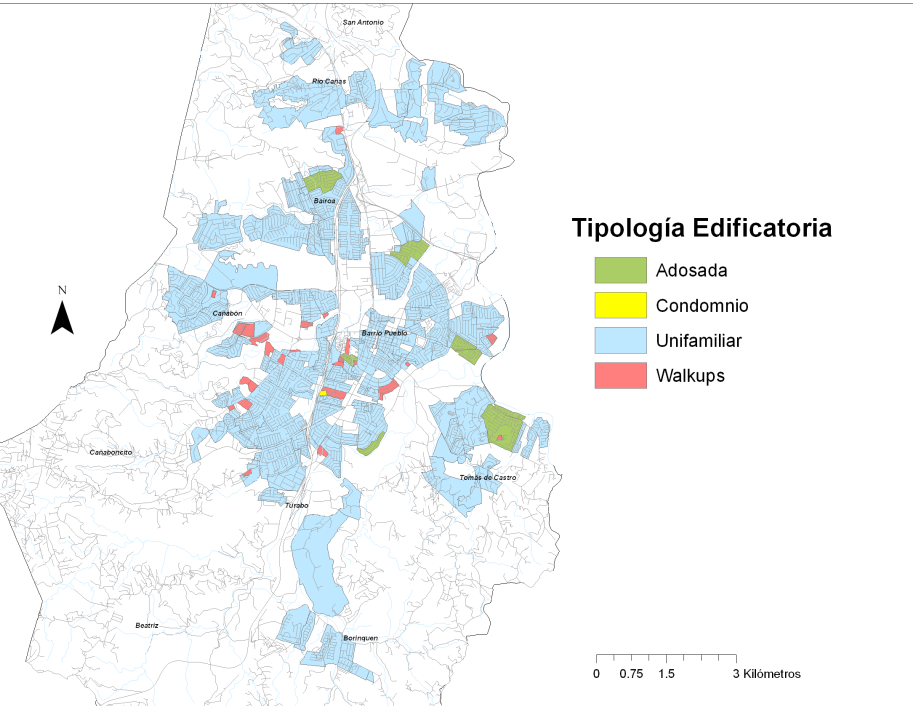
Fuente: Raúl Matos Flores (2016)

Mapa 4.1: Edad de las comunidades que constituyen la zona de estudio



Fuente: Elaboración Propia

Mapa 4.2: Tipo de edificación por comunidad en la zona de estudio



Fuente: Elaboración Propia

Capítulo 5

Análisis Exploratorio de Datos para la Predicción de Valores Inmuebles

5. Análisis Exploratorio de Datos para la Predicción de Valores Inmuebles

Los criterios que se utilizan para la valoración y cobro de contribuciones sobre los inmuebles en Puerto Rico no permiten que el sistema sea justo ni equitativo ya que estos datan de un procedimiento establecido en 1951 y ajustado en 1957. Dado que no se realiza un proceso de actualización de los datos y los dueños de inmuebles no informan de cambios estructurales en términos de mejoras (ampliaciones, terrazas, piscinas, etc.) que han sufrido los inmuebles, estos cambios estructurales del inmueble no están contemplados en el valor de tasación asignado. Se tributa basado en el valor asignado originalmente al inmueble una vez se inscribió y se tasó por primera vez (con las tasas establecidas en 1957).

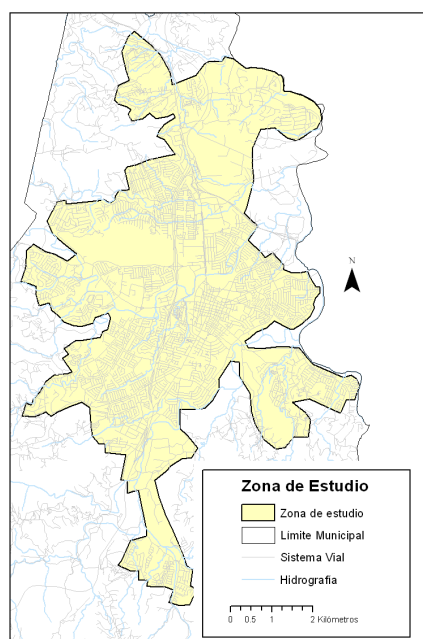
Por lo tanto, el problema que se observa es la ausencia de un proceso de actualización de los valores de inmuebles en Puerto Rico por parte de la agencia encargada. A pesar que las leyes creadas para establecer los procesos de valoración de inmuebles dan espacio para que esta actualización ocurriera, la realidad es que nunca se realizó la misma. El escenario hoy día hace que la tarea de valoración sea más compleja. Por ejemplo, la cantidad de unidades de vivienda en Puerto Rico en el 1950 totalizaban 453.572. El Censo de Población y Vivienda de 2010 indicó que esta cifra llega a 1.636.946. Por lo tanto, los procesos de valoración y actualización se triplican.

Para abordar esta situación se ha diseñado un método de dos etapas. En la primera se busca identificar la forma que se pueda modelar la distribución espacial de valores de inmuebles utilizando las herramientas contemporáneas de análisis. Para esto se compara la técnica de valoración hedónica utilizando Mínimos Cuadrados Ordinarios y la técnica de Regresión Geográficamente Ponderada. La primera se fundamenta en la idea de que, si un bien está constituido por una serie de atributos, se puede asumir que el precio del mercado del bien surge de la suma de los precios individuales de esos atributos. En esta valoración no solo se consideran las características estructurales que constituyen el inmueble, sino también del valor de las externalidades asociadas al mismo. Por otra parte, la técnica de Regresión Geográficamente Ponderada es una técnica de regresión espacial que proporciona un modelo local de la variable o proceso que intenta entender al ajustar una ecuación de regresión a cada entidad en el grupo de datos utilizados. Ambas técnicas serán modeladas dentro de un ambiente SIG y se seleccionará la que mejor desempeño muestre de las dos. Se pusieron a prueba los modelos con datos publicados mensualmente que parten del año 2005 y culminan

en el 2009. Luego, los resultados fueron comparados con valores residenciales de ventas realizadas entre los años 2010 al 2014. Los datos fueron seleccionados para la zona urbana del municipio de Caguas.

Al final se busca que este modelo resuelva el problema de contar con una valoración inmueble actualizada que se pueda aplicar en un tiempo corto cuando es comparado con el posible ejercicio de visitar cada inmueble para tasarlo. El modelo depende de la acción del mismo mercado inmobiliario, porque depende del listado de ventas inmuebles ocurridas en un período de tiempo reciente. Además, cuenta con las características de los inmuebles. Con el período de comprobación (2010-2014) se busca establecer la validez temporal de los valores de inmuebles residenciales como referencia para el establecimiento del componente contributivo del inmueble. Tomar los datos de venta entre los años 2005 y 2009 permite además insertar en el período de análisis el estallido de la burbuja inmobiliaria en el mercado puertorriqueño. Con esto se busca incluir dentro del proceso de validación una condición particular que impacta directamente el mercado inmobiliario y por lo tanto el patrón de los valores inmuebles. Se podría así, no solo evaluar por cuánto tiempo se mantendría la validez de los valores inmuebles para utilizarlos para el proceso contributivo, además se podría observar el impacto en esa validez de un evento singular como es el estallido de la burbuja inmobiliaria.

Gráfica 5.1: Mapa demarcando la zona de estudio seleccionada



5.1 Fuente de los datos

La aplicación de un modelo hedónico obliga a la identificación de una serie de atributos importantes para la predicción del precio del inmueble. Esto crea la necesidad de recopilar datos que permitan ejecutar el modelo y obtener la predicción de valor buscado. En el caso específico de los valores de inmuebles, la especificación hedónica asume que los atributos seleccionados representan características físicas del inmueble, características relacionadas al vecindario y características relacionadas a la localización del inmueble.

Para la etapa del modelo que compara la técnica de valoración hedónica y la técnica de Regresión Geográficamente Ponderada se incluyó inicialmente un total de 2.313 ventas de inmuebles que ocurrieron en la zona de estudio entre los años 2005 y 2009. Se seleccionó este período de manera que se pudiera modelar la situación del mercado de ventas de inmuebles residenciales por el que atravesó Puerto Rico, en particular reflejar el estallido en el año 2007 de la burbuja inmobiliaria por la que atravesaba la isla (Krugman, 2009; Shiller, 2011).

Los datos se obtuvieron de listados producidos por firmas especializadas en generar listados de ventas de inmuebles que mayormente son utilizadas por tasadores en sus funciones de valoración. Los datos para producir estos listados surgen de los informes mensuales producidos por las entidades bancarias que generaron los préstamos hipotecarios. Estos listados de datos no solo incluyen una serie de variables que fueron consideradas por el tasador al momento de generar una opinión de valor sobre el inmueble, sino que también incluye datos tan variados como el nombre del tasador que generó la tasación, el estimado de valor, el valor de la venta final, el banco que generó la hipoteca y otros.

Las características obtenidas de los informes son generales para la isla de Puerto Rico. De los listados utilizados se pudo identificar un total de 119 variables que luego de analizarlas se procedió a agrupar en cinco categorías:

- Condición: Presenta una opinión del tasador sobre la condición de la vivienda al momento de realizar la inspección del inmueble. Esta variable es subjetiva, ya que responde a la impresión del tasador sobre la condición del inmueble. Entre los términos utilizados se encuentran; Excelente, Buena o Pobre.
- Construcción: Establece el material de construcción del inmueble. Dado los códigos de construcción la mayoría de las viviendas se construyen en concreto, pero puede haber casas construidas en madera u otro material.

- Tipo de Venta: Identifica que tipo de préstamo hipotecario fue el aprobado para venta. Los préstamos van desde el de tipo convencional hasta los subvencionados por la “Federal Housing Administration” (para personas que compran su vivienda principal) o los de la Administración de Veteranos.
- Tipología: Establece la tipología del inmueble.
- Otros: Esta categoría agrupa las características estructurales del inmueble (cantidad de baños, cuantas plantas tiene, cantidad de estacionamientos, cantidad de dormitorios) y las características de la comunidad donde se ubica (piscina, cancha de baloncesto, campo de Golf).

Para la selección de las variables a incluir en el modelo se tabularon las 88 variables relacionadas a las características estructurales y comunitarias de los inmuebles. Se creó un listado con todas las variables identificadas en la base de datos y luego se procedió a contabilizar la cantidad de inmuebles que contaba con cada una de ellas. En cuanto a las variables estructurales del inmueble, se seleccionaron aquellas que se entiende son las más importantes al momento de influenciar en una venta. Por ejemplo, la superficie del solar, la superficie de ocupación, la edad del inmueble, la cantidad de baños o la cantidad de dormitorios. Sobre las características comunitarias, todas aquellas que no se manifestaban en la zona de estudio fueron eliminadas. Aquellas que ocurrían en menos de un 25 por ciento de los casos fueron analizadas una a una y las que se consideraban importantes se incluían, las otras se excluían. El restante de las variables fue incluido.

Además de las variables obtenidas de los listados se buscó añadir variables relacionadas a aspectos espaciales del inmueble, por ejemplo, la distancia espacial para la obtención de algún tipo de servicio. Gran parte de la literatura revisada considera la distancia al centro de negocios o distancia a escuelas como factores imprescindibles al momento de analizar valores residenciales por medio de modelo hedónico. El problema es que estas condiciones responden a la realidad funcional de las ciudades en los Estados Unidos y otros países. En Puerto Rico las escuelas del estado no regulan el servicio a niños en función del lugar de residencia. Las escuelas privadas se rigen por el libre mercado y aceptan a cualquier niño siempre que cumpla con los requisitos de admisión y la familia pueda costear los estudios. El prestigio que una escuela privada puede generar que su ámbito de mercado se amplíe a distancias que sobrepasan en ocasiones varios municipios. Por lo tanto, el factor de cercanía geográfica a una escuela tiene muy poca influencia al momento de decidir comprar un

inmueble. Dado que la mayoría de los nuevos desarrollos son para personas de altos recursos económicos, estos son mercado de las escuelas privadas del más alto prestigio. Por tal razón se analizó qué variables relacionadas con distancia a servicios son importantes al momento de adquirir un inmueble. Se entiende que la distancia al centro comercial más cercano y la distancia a un conector del sistema de autovías del país son factores que influyen en la venta. Las promociones de ventas de hogares destacan estos dos factores al momento de presentar los nuevos complejos residenciales. Utilizando la herramienta de análisis de proximidad se calculó la distancia de cada residencia al centro comercial más cercano y al conector del sistema de autovías más cercano.

Luego de aplicado el proceso de depuración se seleccionaron las variables que se someterían al diagnóstico preliminar para determinar cuáles finalmente se utilizarían en el modelo. Estas fueron:

- Variable dependiente
 - Saleprice: Precio de venta del inmueble en dólares.
- Variables Independientes
 - Estructurales de la residencia
 - Lotarea: Superficie del solar medido en metros cuadrados.
 - Age: Años que lleva construido el inmueble.
 - Bedrooms: Cantidad de dormitorios con que cuenta el inmueble.
 - Bathrooms: Cantidad de baños con que cuenta el inmueble.
 - HouseareaM: Superficie de ocupación del inmueble medido en metros cuadrados.
 - Comunitarios
 - ACC: Si la comunidad es una cerrada y cuenta con control de acceso.
 - GZB: Si las facilidades recreativas de la comunidad cuentan con un gazebo⁷ para actividades.
 - PG: Existencia de área de juego para niños.
 - SWP: Existencia de piscina.
 - TENIS: Existencia de canchas de tenis.

⁷ Un gazebo es un pabellón de planta simétrica, generalmente hexagonal o circular, que comúnmente se encuentra en los parques, jardines, y en áreas públicas abiertas. Los gazebos se encuentran aislados, techados, y abiertos por todos los lados; proporcionan sombra, un abrigo de características básicas, función ornamental en un paisaje, y un lugar de descanso.

- BSKT: Existencia de canchas de baloncesto.
- PP: Existencia de un parque pasivo.
- GYM: Existencia de gimnasio.
- CH: Si el complejo residencial tiene una Casa Club para reuniones.
- VB: Existencia de cancha de voleibol.
- BB: Existencia de parque de béisbol.
- GLF: Existencia de campo de golf.
- Distancia a servicios (Localización)
 - CC_DIST: Distancia euclidiana a centro comercial más cercano.
 - EXIT_DIST: Distancia euclidiana a conector de vía de circulación nacional más cercano.

Estas variables fueron sometidas a un proceso de diagnóstico para determinar finalmente cuales serían utilizadas en el modelo hedónico. Previo a este diagnóstico los valores de ventas fueron deflactados utilizando los datos del Índice de Precios al Consumidor publicados por el Departamento del Trabajo de Puerto Rico.

Al observar los valores de estos inmuebles se puede destacar lo siguiente: (1) la venta con el precio más bajo ocurrió en el año 2006, 49.000 dólares (2) en ese mismo año ocurrió la venta del precio más alto, 846.000 dólares, (3) el promedio de tamaño de los solares ronda los 350 m² y el predio con mayor tamaño fue de 13.210 m² y (4) el inmueble con mayor superficie de ocupación fue una residencia de 368 m².

En términos de variables comunitarias, se contabilizaron un total de 79 sectores o comunidades con control de acceso, 45 contaban con piscina, 52 con gazebos, 56 con área de juego para niños, 23 con canchas de tenis, 50 con canchas de baloncesto y una con campo de golf.

Tabla 5.1: Estadísticas Descriptivas para variables en el 2005 (Total de Ventas: 511)

Variable	Descripción	Min	Max	Media	Desv. Est.
Saleprice	Venta inmueble \$	62.000	825.000	206.765	126.405
Lotarea	Solar (m ²)	0	11.791	365	623
Age	Edad propiedad (años)	0	60	10	12
Bedrooms	Dormitorios	1	8	3	0,79
Bathrooms	Baños	1	4.5	2	0,88
HouseareaM	Ocupación (m ²)	54	351	149,60	71,38
ACC	Control acceso	0	1	0,68	0,46
GZB	Mirador (Gazebo)	0	1	0,60	0,48
PG	Área de Juego	0	1	0,66	0,47
SWP	Piscina	0	1	0,56	0,49
TENIS	Cancha de Tenis	0	1	0,42	0,49
BSKT	Baloncesto	0	1	0,40	0,49
PP	Parque Pasivo	0	1	0,18	0,38
GYM	Gimnasio	0	1	0,37	0,48
CH	Casa Club	0	1	0,41	0,49
VB	Cancha Voleibol	0	1	0,16	0,37
BB	Baseball	0	1	0,23	0,42
GLF	Campo de Golf	0	1	0,005	0,07
CC_DIST	Dist. Cen comercial (m)	154	5790	1.478	1.097
EXIT_DIST	Dist Autopista (m)	97	5978	1.875	1.046

Tabla 5.2: Estadísticas Descriptivas para variables en el 2006 (Total de Ventas: 867)

Variable	Descripción	Min	Max	Media	Desv. Est.
Saleprice	Venta inmueble \$	49.000	846.000	198.200	134.911
Lotarea	Solar (m ²)	0	13.210	398,48	646,19
Age	Edad propiedad (años)	0	55	13	13
Bedrooms	Dormitorios	2	9	3	0,73
Bathrooms	Baños	1	5	2	0,76
HouseareaM	Ocupación (m ²)	55	355	134,71	63,39
ACC	Control acceso	0	1	0,59	0,49
GZB	Mirador (Gazebo)	0	1	0,50	0,49
PG	Área de Juego	0	1	0,54	0,49
SWP	Piscina	0	1	0,46	0,49
TENIS	Cancha de Tenis	0	1	0,30	0,46
BSKT	Baloncesto	0	1	0,37	0,48
PP	Parque Pasivo	0	1	0,11	0,31
GYM	Gimnasio	0	1	0,21	0,41
CH	Casa Club	0	1	0,25	0,43
VB	Cancha Voleibol	0	1	0,09	0,29
BB	Baseball	0	1	0,20	0,40
GLF	Campo de Golf	0	1	0,01	0,11
CC_DIST	Dist. Cen comercial (m)	165	6117	1657	1210
EXIT_DIST	Dist Autopista (m)	105	6275	1798	1053

Tabla 5.3: Estadísticas Descriptivas para variables en el 2007 (Total de Ventas: 226)

Variable	Descripción	Min	Max	Media	Desv. Est.
Saleprice	Venta inmueble \$	53.000	699.500	177.893	122.909
Lotarea	Solar (m ²)	0	1.764	325,93	263,61
Age	Edad propiedad (años)	0	47	17	13
Bedrooms	Dormitorios	2	7	3	0,68
Bathrooms	Baños	1	4.5	1	0,73
HouseareaM	Ocupación (m ²)	56	357	123,37	58,34
ACC	Control acceso	0	1	0,46	0,49
GZB	Mirador (Gazebo)	0	1	0,33	0,47
PG	Área de Juego	0	1	0,43	0,49
SWP	Piscina	0	1	0,26	0,44
TENIS	Cancha de Tenis	0	1	0,18	0,38
BSKT	Baloncesto	0	1	0,36	0,48
PP	Parque Pasivo	0	1	0,09	0,29
GYM	Gimnasio	0	1	0,12	0,32
CH	Casa Club	0	1	0,17	0,37
VB	Cancha Voleibol	0	1	0,09	0,29
BB	Baseball	0	1	0,21	0,41
GLF	Campo de Golf	0	1	0,004	0,06
CC_DIST	Dist. Cen comercial (m)	140	5.206	1.592	1.141
EXIT_DIST	Dist Autopista (m)	161	3.812	1.671	1.072

Tabla 5.4: Estadísticas Descriptivas para variables en el 2008 (Total de Ventas: 466)

Variable	Descripción	Min	Max	Media	Desv. Est.
Saleprice	Venta inmueble \$	59.900	656.973	196.120	109.799
Lotarea	Solar (m ²)	0	5524	330,74	436,35
Age	Edad propiedad (años)	0	60	13	14
Bedrooms	Dormitorios	0	7	3	0,73
Bathrooms	Baños	1	4.5	2	0,72
HouseareaM	Ocupación (m ²)	59	368	132,74	56,61
ACC	Control acceso	0	1	0,53	0,49
GZB	Mirador (Gazebo)	0	1	0,45	0,49
PG	Área de Juego	0	1	0,48	0,49
SWP	Piscina	0	1	0,42	0,49
TENIS	Cancha de Tenis	0	1	0,30	0,46
BSKT	Baloncesto	0	1	0,33	0,47
PP	Parque Pasivo	0	1	0,09	0,29
GYM	Gimnasio	0	1	0,27	0,44
CH	Casa Club	0	1	0,29	0,45
VB	Cancha Voleibol	0	1	0,14	0,35
BB	Baseball	0	1	0,17	0,38
GLF	Campo de Golf	0	1	0,01	0,12
CC_DIST	Dist. Cen comercial (m)	140	5.589	1.645	1.125
EXIT_DIST	Dist Autopista (m)	103	5.766	1.843	1.037

Tabla 5.5: Estadísticas Descriptivas para variables en el 2009 (Total de Ventas: 229)

Variable	Descripción	Min	Max	Media	Desv. Est.
Saleprice	Venta inmueble \$	50.000	627.803	171.204	98.112
Lotarea	Solar (m ²)	0	8.012,51	405,55	616,03
Age	Edad propiedad (años)	0	58	19	15
Bedrooms	Dormitorios	2	8	3	0,65
Bathrooms	Baños	1	4,5	2	0,74
HouseareaM	Ocupación (m ²)	43	348	126,31	56,90
ACC	Control acceso	0	1	0,36	0,48
GZB	Mirador (Gazebo)	0	1	0,31	0,46
PG	Área de Juego	0	1	0,31	0,46
SWP	Piscina	0	1	0,27	0,44
TENIS	Cancha de Tenis	0	1	0,17	0,37
BSKT	Baloncesto	0	1	0,30	0,45
PP	Parque Pasivo	0	1	0,08	0,27
GYM	Gimnasio	0	1	0,12	0,32
CH	Casa Club	0	1	0,12	0,33
VB	Cancha Voleibol	0	1	0,82	0,27
BB	Baseball	0	1	0,22	0,41
GLF	Campo de Golf	0	1	0,004	0,06
CC_DIST	Dist. Cen comercial (m)	148	5.298	1.672	1.033
EXIT_DIST	Dist Autopista (m)	116	5.303	1.681	1.032

5.2 Diagnóstico del modelo de regresión inicial

Antes de adoptar las relaciones establecidas en el modelo y aceptarlas como correctas es importante analizar el desempeño de éstas. Para esto se realizaron una serie de pruebas estadísticas que ayudaron a determinar el desempeño general del modelo. Las primeras pruebas realizadas fueron el Coeficiente de determinación múltiple R^2 y la del R^2 ajustado. Estas pruebas indican cuanta variación de la variable dependiente es explicada por el modelo (Rosenshein, et al., 2011). La prueba arroja valores que van del 0,0 a 0,1. Al analizar los datos se debe recordar que el R^2 Ajustado siempre tiende a mostrar valores menores que el Coeficiente de determinación múltiple, ya que muestra la complejidad del modelo. Esta es una medida más precisa del desempeño del modelo. El valor obtenido del coeficiente indica el porcentaje de variación en la variable dependiente que es explicado por el modelo.

Los modelos para el período bajo estudio (2005 – 2009) mostraron valores R^2 Ajustado que fluctúan entre 0,78 y 0,90 (Tabla 5.6). En otras palabras, en este período de tiempo las variables independientes seleccionadas pueden explicar entre un 78% y un 90% el comportamiento de la variable dependiente. Al comparar con estudios similares, estos valores muestran un desempeño excepcional. Por ejemplo, el resultado de Orford (1999) de un 84%

se considera de muy buen desempeño. En otros estudios los valores fueron superados significativamente. Como ejemplo se puede citar el 62% obtenido por Yu (2007) en su estudio en Milwaukee. Por tal razón se concluye que es acertada la selección de las variables explicativas escogidas dado el porcentaje tan alto obtenido en esta prueba.

Tabla 5.6: Coeficiente de determinación múltiple y R^2 ajustado por año.

Prueba	2005	2006	2007	2008	2009
Coeficiente determinación múltiple	0,953	0,924	0,937	0,929	0,898
R^2 ajustado	0,905	0,851	0,878	0,857	0,789

En el proceso de diagnóstico también es importante evaluar las variables explicativas utilizadas en el modelo. Para esto se utiliza el resultado de los coeficientes (Tabla 5.7). Cada variable explicativa refleja la fortaleza y el tipo de relación que tienen las variables explicativas con la variable dependiente. Cuando el signo asociado al coeficiente es negativo, la relación es negativa. Cuando el signo es positivo, la relación es positiva (Rosenshein, et al., 2011). La validación de los resultados de los coeficientes comparándolos con otros estudios se torna dificultosa debido a dos razones principales (Sirmans et al., 2005). La primera de éstas, los valores de venta reflejan la predilección de los compradores por características particulares del inmueble. Esto influye en el resultado del coeficiente. La otra razón es que las variables utilizadas en los estudios no se describen dentro de la base de datos de forma uniforme. Por estas razones se sugiere que los resultados obtenidos de esta prueba solo sean analizados con relación a lo esperado con el signo del coeficiente y el resultado final obtenido (Rosenshein, et al., 2011).

Para analizar este aspecto se construyó una tabla (Tabla 5.8) indicando si el valor esperado para cada variable explicativa era positivo o negativo y se comparó con el resultado obtenido para los 5 años de estudio. Se esperaba que 16 variables obtuvieran un coeficiente positivo y tres un coeficiente negativo. En términos de variables explicativas de tipo positivo solo la superficie de la parcela (LOTAREA), la superficie de ocupación (HOUSEAREA), la piscina (SWP), la cancha de volibol (VB) y el parque de béisbol (BB) presentaron coeficientes esperados en los cinco años. Por su parte, de las variables explicativas con efecto negativo, solo la edad de la propiedad (AGE) y la distancia al centro comercial más cercano

(CC_DIST) arrojaron coeficientes negativos en los cinco años. Las restantes variables mostraron coeficientes que variaban entre positivos y negativos en los diversos años. Las que más resaltan son la de cantidad de baños (BATHROOMS) y cantidad dormitorios (BEDROOMS). En el caso de Puerto Rico existe la idea que el comprador de residencias busca la mayor comodidad y dentro de este concepto se busca la mayor cantidad de baños y dormitorios posible. En el caso de baños (BATHROOMS) se esperaba que a mayor cantidad de estos, el valor del inmueble aumentara y el signo negativo del coeficiente presenta lo contrario. Tres de los cinco años reflejaron coeficientes negativos. El caso de los dormitorios (BEDROOMS) es más significativo. Domina en el mercado de los gestores inmobiliarios en Puerto Rico la idea que el comprador de inmuebles busca dormitorios adicionales en sus hogares con el objetivo de utilizar estos para oficinas personales, bibliotecas, salones de juego para los niños o una habitación adicional para huéspedes. Se esperaba que una mayor cantidad de dormitorios significara un valor mayor en el inmueble. El resultado da la impresión de lo contrario ya que en los cinco años de estudio el coeficiente resultó ser negativo. Estos coeficientes opuestos a lo esperado pueden ser efecto de confusión en la interpretación del signo (Kennedy, 2005). En su interpretación sobre este fenómeno Kennedy (2005) establece que el coeficiente “baño” indica que, a mayor cantidad de estos, el valor del inmueble aumenta, mientras que otros regresores como puede ser la superficie de ocupación se mantiene constante. Por lo tanto, la adición de un baño bajo esta restricción debe implicar una reducción en metros cuadrados en otras habitaciones del inmueble, como dormitorios más pequeños o la pérdida de un comedor, lo que conllevará una pérdida de valor del inmueble. En este caso, el efecto neto sobre el precio es negativo. Para estimar el impacto en el precio del inmueble de la adición de un baño, por ejemplo, hay que tener en cuenta las contribuciones tanto del propio baño como del aumento en metros cuadrados del mismo (Kennedy, 2005).

Por otro lado, en el caso de la distancia a la conexión más cercana del sistema de autopistas se esperaba que los resultados fueran negativos y los coeficientes resultaron todos positivos. Es importante establecer que los coeficientes son dados en la misma unidad de sus variables explicativas asociadas. Estos reflejan el cambio esperado en la variable dependiente por cada unidad de cambio de la variable explicativa.

Tabla 5.7: Coeficientes

Variable	2005	2006	2007	2008	2009
LOTAREA	12,880	14,463	179,206	15,818	12,058
AGE	-743,933	-997,704	-1834,109	-680,312	-1678,030
BEDROOMS	-11214,689	-15174,541	-16829,034	-5122,484	-7573,531
BATHROOMS	-22598,268	-10075,266	7098,792	-1987,559	1599,481
CC_DIST	-3,551	-4,291	-9,595	-12,737	-16,771
EXIT_DIST	9,844	5,238	6,461	22,180	18,878
HOUSEAREAM	1754,388	1733,440	907,760	1361,767	1078,103
ACC	24692,994	23781,189	-6106,840	17876,672	-2321,585
GZB	22319,611	-1854,700	19776,037	-159,572	8899,318
PG	-9578,079	-16875,125	14002,151	-6564,841	-21944,561
SWP	1294,713	16029,247	8626,653	29199,521	19389,739
TENIS	10374,239	12103,913	6494,622	-7982,089	5819,904
BSKT	-9,513	-487,156	-9361,644	247,815	-5496,508
PP	-20580,584	-10898,693	-6534,175	-9812,479	-25890,101
GYM	-56801,212	-32691,300	30342,656	-8674,429	-28565,089
CH	29752,641	6314,711	-6329,233	-23024,494	7696,061
VB	27255,223	48730,391	45620,123	9166,148	43071,069
BB	33791,899	28600,148	7852,576	21930,788	7413,230
GLF	262559,956	99676,530	32903,337	32046,646	-9579,844

Tabla 5.8: Relaciones Esperadas y obtenidas para cada año (Positivas / Negativas)

Variable	Esperada	2005	2006	2007	2008	2009
LOTAREA	P	P	P	P	P	P
AGE	N	N	N	N	N	N
BEDROOMS	P	N	N	N	N	N
BATHROOMS	P	N	N	P	N	P
CC_DIST	N	N	N	N	N	N
EXIT_DIST	N	P	P	P	P	P
HOUSEAREAM	P	P	P	P	P	P
ACC	P	P	P	N	P	N
GZB	P	P	N	P	N	P
PG	P	N	N	P	N	N
SWP	P	P	P	P	P	P
TENIS	P	P	P	P	N	P
BSKT	P	N	N	N	P	N
PP	P	N	N	N	N	N
GYM	P	N	N	P	N	N
CH	P	P	P	N	N	P
VB	P	P	P	P	P	P
BB	P	P	P	P	P	P
GLF	P	P	P	P	P	N

Otra prueba utilizada para evaluar el desempeño del modelo fue el de la Probabilidad y la Probabilidad Robusta. Estas miden la significancia estadística del coeficiente (Rosenshein et al., 2011). En el informe de resultados producido se coloca un asterisco a las probabilidades estadísticamente significativas. Una variable explicativa asociada a un coeficiente estadísticamente significativo es importante para el modelo de regresión, si el sentido común asume una relación válida con la variable dependiente.

Los resultados para los cinco años bajo estudio reflejan que algunas variables se demarcan como significativas de manera regular a lo largo del período (Tabla 5.9). Estas son el tamaño del lote (LOTAREA), la edad del inmueble (AGE), y la superficie de ocupación del inmueble (HOUSEAREAM). Estas concuerdan con lo planteado por Sirmans et al. (2005) cuando analizan las variables más comunes en procesos de valoración hedónica de inmuebles residenciales. En otro grupo se identifican variables con probabilidades mayores a 0,05% en los cinco años analizados. Un tercer grupo de variables explicativas muestran resultados menores al 0,05% en el período de estudio. Las variables con probabilidades mayores a 0,05% en el período de estudio fueron la existencia de cancha de tenis (TENIS) y la existencia de cancha de baloncesto (BSKT). Las restantes variables se pueden dividir entre aquellas que en tres años o más mostraron valores significativos como son el número de dormitorios (BEDROOMS) (en tres de los años), la existencia de control de acceso para la comunidad (ACC) (en tres de los años), la existencia de cancha de voleibol (VB) (cuatro de los años) y el parque de béisbol (tres de los años). Esto resulta interesante ya que en estudios previos estas variables no suelen ser significativas en los modelos de valoración hedónica para inmuebles (Sirmans et al., 2005).

Por otro lado, el Factor de Inflación de la Varianza (FIV) mide la redundancia entre las variables explicativas (Burt et al., 2009). En otras palabras, ayuda a detectar la multicolinealidad. Cuanto mayor sea la relación entre las variables independientes más variarán los coeficientes, por lo que será mayor el FIV. Como regla general las variables explicativas con valores FIV superiores a 7,5 deben ser eliminadas una a una del modelo de regresión. Las variables que mostraron esta condición en más de una ocasión (Tabla 5.10) fueron la existencia de gazebo para actividades (GZB), la existencia de gimnasio (GYM) y de casa club (CH). Con valores FIV más bajos de manera constante en los cinco años

analizados se puede mencionar el tamaño del lote (LOTAREA), la edad del inmueble (AGE), el parque de béisbol (BB) y el campo de golf (GLF).

Tabla 5.9: Variables Significativas (Probabilidad menor a 0,05 (95% de confianza))

Variable	2005	2006	2007	2008	2009
LOTAREA	0,000455*	0,000011*	0,000000*	0,002424*	0,024295*
AGE	0,030554*	0,000064*	0,000002*	0,000308*	0,000000*
BEDROOMS	0,000746*	0,000012*	0,005344*	0,066619	0,204647
BATHROOMS	0,000007*	0,036037*	0,367677	0,654373	0,824750
CC_DIST	0,341438	0,154082	0,107862	0,000116*	0,002009*
EXIT_DIST	0,021225*	0,131719	0,322450	0,000000*	0,001156*
HOUSEAREAM	0,000000*	0,000000*	0,000000*	0,000000*	0,000000*
ACC	0,013718*	0,003836*	0,632594	0,021062*	0,847730
GZB	0,048309*	0,852643	0,176624	0,986894	0,552296
PG	0,289023	0,043992*	0,219709	0,418738	0,170554
SWP	0,895717	0,079624	0,595824	0,007485*	0,220116
TENIS	0,164515	0,064558	0,643698	0,277438	0,623345
BSKT	0,998673	0,926975	0,304153	0,965262	0,587329
PP	0,009721*	0,139896	0,628480	0,299490	0,075508
GYM	0,000050*	0,004452*	0,181733	0,508564	0,321978
CH	0,012119*	0,524523	0,692130	0,055087	0,780947
VB	0,004605*	0,000001*	0,002781*	0,276923	0,007338*
BB	0,000001*	0,000003*	0,379750	0,000649*	0,384203
GLF	0,000000*	0,000001*	0,514657	0,101003	0,835868

TABLA 5.10: VIF – Factor de Inflación de Varianza (7,5)

Variable	2005	2006	2007	2008	2009
LOTAREA	1,430	1,387	2,266	1,664	1,427
AGE	5,373	3,529	2,837	2,437	2,675
BEDROOMS	1,902	2,001	1,965	1,372	2,045
BATHROOMS	5,197	4,278	3,885	3,400	3,835
CC_DIST	4,643	4,233	5,386	4,372	4,088
EXIT_DIST	5,505	4,274	5,720	4,073	4,647
HOUSEAREAM	8,590	5,837	5,941	5,666	5,365
ACC	5,966	5,179	4,739	4,825	4,510
GZB	8,443	7,952	5,524	7,600	6,414
PG	5,058	5,541	3,721	5,342	7,308
SWP	6,650	6,632	6,085	9,373	6,534
TENIS	3,767	2,917	3,483	3,741	2,687
BSKT	2,183	2,119	2,236	2,351	2,866
PP	2,639	1,726	1,795	2,481	2,128
GYM	12,383	7,124	6,518	11,082	11,831
CH	9,389	5,931	4,260	9,632	11,249
VB	3,525	2,527	2,240	2,873	2,564
BB	2,220	1,839	1,583	1,934	1,689
GLF	1,347	1,565	1,311	1,829	1,235

Otro indicativo para determinar la fiabilidad en los modelos es la distribución de los valores residuales. En un modelo propiamente especificado estos deben contar con una distribución normal y una media de cero (Burt et al., 2009). Para esto se utilizó la prueba Jarque – Bera (JB). Esta prueba establece si los residuos del modelo de regresión están normalmente distribuidos. En otras palabras, la prueba Jarque – Bera (JB) busca establecer cuánto se desvían los coeficientes de asimetría y curtosis.

La prueba Jarque – Bera utiliza tanto el valor de la asimetría como el de la curtosis para su análisis. La fórmula de Jarque – Bera es:

dónde:

$$JB = \frac{T-k}{6} \left[S^2 + \frac{(K-3)^2}{4} \right]$$

- T: Tamaño de muestra
- K: Es la kurtosis
- S: Es la asimetría

La hipótesis nula para esta prueba es que los residuales son distribuidos normalmente, por lo tanto, si se construyera un histograma de los residuales, éste se asemejaría a una curva de Bell o una distribución gaussiana. Cuando el valor p (probabilidad) de esta prueba es bajo (menor que 0,05 para un nivel de confianza del 95 por ciento, por ejemplo), los residuales no son distribuidos normalmente, lo que indica que el modelo es tendencioso. Si también cuenta con una autocorrelación espacial estadísticamente significativa de los residuales, la influencia puede ser el resultado de una especificación incorrecta del modelo, por ejemplo, que una variable clave está faltando en el modelo. Los resultados de un modelo de MCO con especificación incorrecta no son confiables. También puede producirse una prueba Jarque Bera estadísticamente significativa si se intenta modelar relaciones no lineales, si los datos incluyen valores atípicos influyentes o cuando hay una heterocedasticidad.

La regla de decisión es que si JB es menor a 5,99 y el valor p está por debajo de 0,05 se rechaza la hipótesis nula, por lo tanto, los residuales del modelo estarían distribuidos de forma normal (Rosenshein et al., 2011). Para los años bajo estudio se realizó esta prueba y los resultados resultaron ser todos superiores a 5,99 junto a valores p por debajo de 0,05 (Tabla 5.11). Con esto se concluye que, en el caso de estudio, ninguno de los períodos analizados cumple con esta regla y por lo tanto los residuos del modelo de regresión no están distribuidos normalmente. El valor obtenido más bajo es el del año 2009 con 41,29. Aunque

bajo, queda muy lejos del valor esperado. El análisis visual de la distribución de los residuales por medio de histogramas (Gráfica 5.3) ayuda a entender mejor esta situación. En ellos se puede observar la existencia de valores atípicos, los cuales pueden influenciar significativamente el resultado de la prueba.

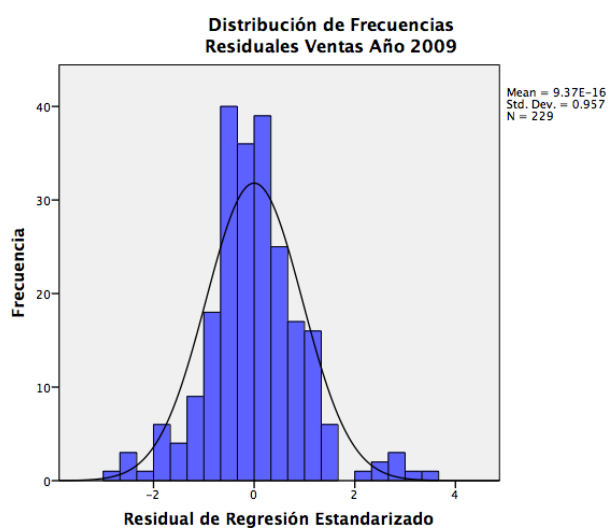
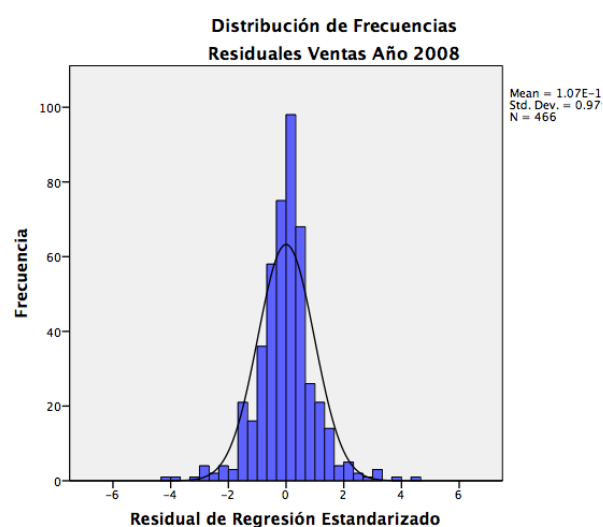
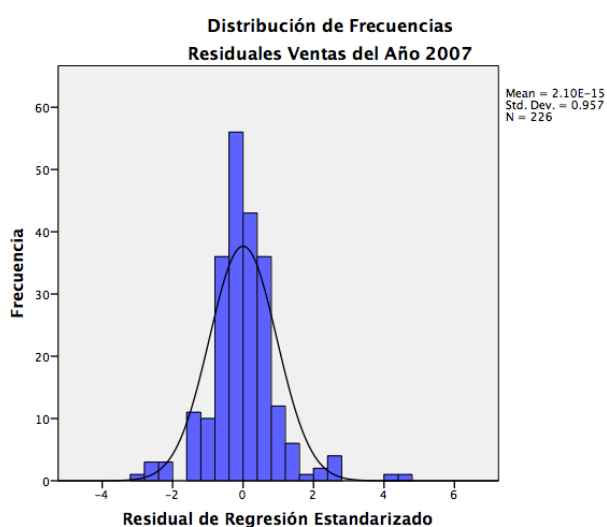
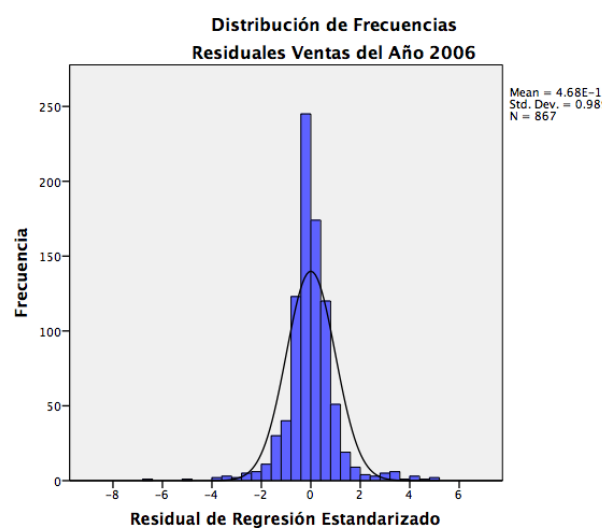
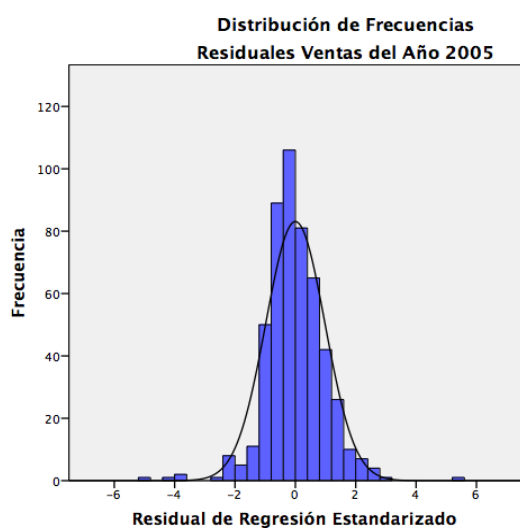
TABLA 5.11: Resultado Jarque - Bera

	2005	2006	2007	2008	2009
Estadístico Jarque-Bera	396,1088	2049,7115	211,6245	194,6131	41,2937
p-Valor	0,0000*	0,0000*	0,0000*	0,0000*	0,0000*

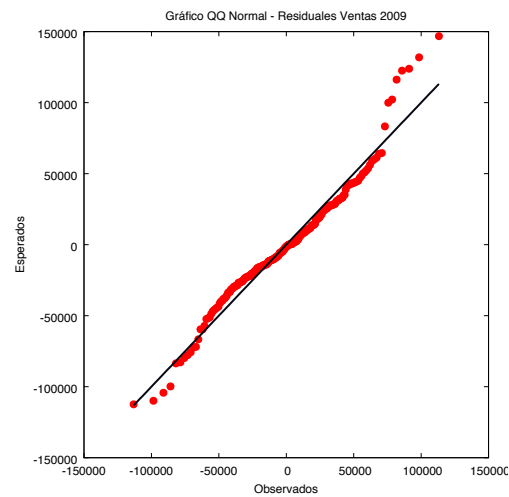
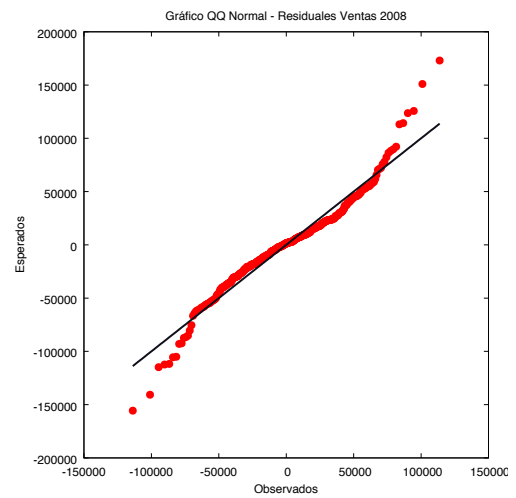
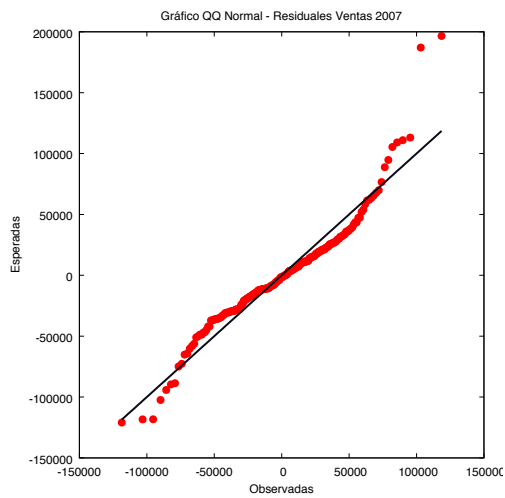
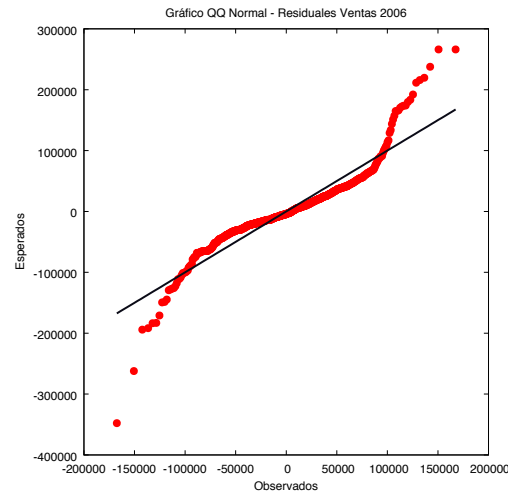
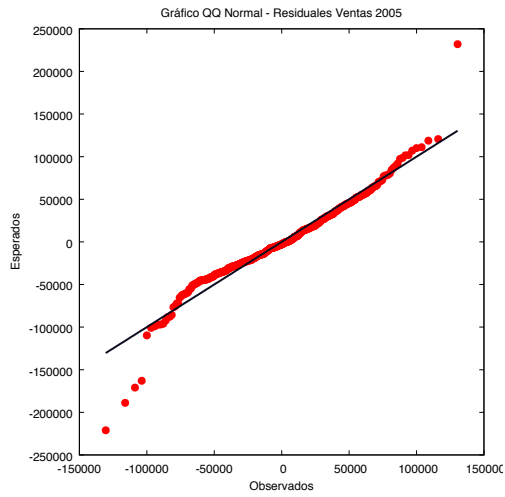
En el proceso de evaluación del modelo también se utilizaron los gráficos Cuantil-Cuantil (Q-Q). Un gráfico Cuantil-Cuantil permite observar cuan cerca está la distribución de un conjunto de datos a alguna distribución ideal o comparar la distribución de dos o más conjuntos de datos. Los gráficos para los años bajo análisis muestran el efecto que tienen valores extremos en alejar la distribución a lo que se espera en cada año. Este efecto se presenta con mayor impacto en valores de ventas más altos.

Excepto los extremos de las distribuciones, se puede observar que los puntos se encuentran todos muy cerca de la línea, lo suficiente como para establecer que estos datos pudieran provenir de una distribución normal. Se puede observar un poco de movimiento alrededor de la línea, sin embargo, esto no debe descalificar a estos datos de estar distribuidos normalmente.

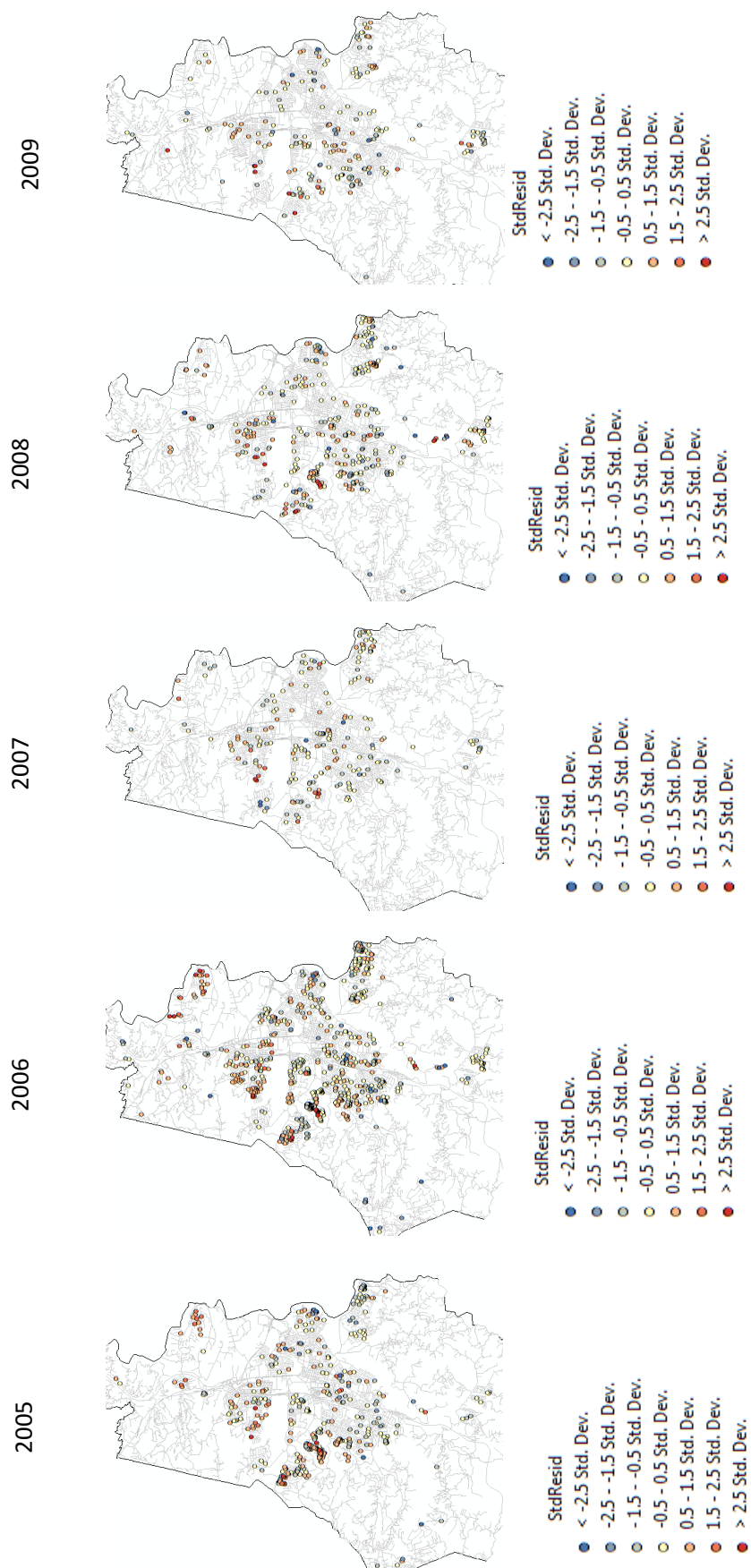
Gráfica 5.2: Histogramas de Residuales, 2005 - 2009



Gráfica 5.3: Gráficos QQ Normal – Residuales Ventas 2005-09



Gráfica 5.4: Mapas de Residuales. Ventas Residenciales 2005 - 2009



Los residuales también pueden ser analizados desde el punto de vista de distribución espacial. Para esto se examina el mapa de residuales. Los patrones de distribución espacial en los residuales del modelo pueden proporcionar pistas acerca de posibles condiciones espaciales que condicionan el comportamiento de estos. En el caso de estudio, para los cinco años, se pudo observar que las zonas de agrupamiento de valores altos se relacionan con los desarrollos urbanos (urbanizaciones) cuya fase final de construcción y comienzo de venta de los inmuebles coincide con los años seleccionados para el estudio.

Otro aspecto considerado en el diagnóstico es la estacionariedad. El índice estadístico de Koenker-BP (Índice estadístico de Breusch-Pagan estudentizado de Koenker) es una prueba utilizada para determinar si las variables explicativas del modelo tienen una relación consistente con la variable dependiente, tanto en el espacio geográfico como en el espacio de datos (Rosenshein et al., 2011). Cuando el modelo es consistente en el espacio geográfico, los procesos espaciales representados por las variables explicativas se comportan de la misma manera en cualquier parte del área de estudio, por lo tanto, los procesos son estacionarios. Por otro lado, cuando el modelo es consistente en el espacio de datos, la variación en la relación entre los valores previstos y cada variable explicativa no cambia cuando cambian las magnitudes de la variable explicativa. En otras palabras, no hay heterocedasticidad en el modelo.

La hipótesis nula de esta prueba es que el modelo es estacionario. Un indicativo para comprobar ésta es observar el resultado del p Valor. Si éste es significativo, la relación no es consistente a lo largo de la zona de estudio y demuestra no estacionariedad. En otras palabras, las relaciones entre las variables varían a lo largo de la zona de estudio. Para un nivel de confianza de 95 por ciento, una probabilidad menor a 0,05 indica una heterocedasticidad significativa y no-estacionariedad.

Al observar los resultados de esta prueba en los cinco años analizados (Tabla 5.12) se muestran valores menores al 0,05. Se asume que los modelos de cada año cuentan con una heterocedasticidad significativa y son no-estacionarios. Los modelos de regresión con no-estacionariedad estadísticamente significativa, generalmente son buenos candidatos para el análisis de Regresión Ponderada Geográficamente (GWR).

TABLA 5.12: Estacionariedad del modelo (Koenker / Breuch-Pagan)

Prueba	2005	2006	2007	2008	2009
Estadística Koenker / Breush – Pagan	135,574	234,149	82,428	185,596	94,409
Valor p	0,000*	0,000*	0,000*	0,000*	0,000*

Finalmente, en este proceso de evaluación del desempeño del modelo se considera el aspecto de la autocorrelación espacial. La autocorrelación espacial refleja el grado en que los objetos o actividades en una unidad geográfica, son similares a otros objetos o actividades en unidades geográficas próximas (Anselin y Bera, 1998). La propiedad básica de los datos espacialmente autocorrelacionados es que los valores no son aleatorios en el espacio, es decir, los valores de los datos están espacialmente relacionados entre sí. La autocorrelación espacial puede presentarse con valores positivos o negativos. Existe autocorrelación positiva cuando valores similares de una variable aleatoria tienden a aglomerarse en el espacio, habiendo dependencia espacial entre ellos; por otra parte, la autocorrelación negativa se presenta cuando las unidades geográficas de observación tienden a estar rodeadas de valores opuestos estadísticamente significativos (Anselin y Bera, 1998). Las aglomeraciones resultantes son caracterizadas por la similitud de los valores de las unidades de observación, clasificándose en altas, con valores superiores a la media del conjunto de observaciones y bajas, con valores inferiores a esa misma media.

Para evaluar la autocorrelación espacial residual se puede calcular el Índice de Moran. Esta prueba mide la autocorrelación espacial basada en la localización de los elementos y su respectivo valor de manera simultánea (Anselin y Bera, 1998). Ésta se considera una estadística deductiva, lo que significa que los resultados del análisis siempre se interpretan dentro del contexto de la hipótesis nula. Para el Índice de Moran global, la hipótesis nula establece que el atributo que se analiza está distribuido en forma aleatoria entre las entidades del área de estudio; es decir, no existe concentración espacial. Un índice cerca de +1,0 indica máxima concentración mientras que un índice cercano a -1,0 representa máxima dispersión. Un valor de 0 representaría un patrón espacial totalmente aleatorio.

El índice de Moran también calcula el valor Z. Éste indica si se rechaza o no la hipótesis nula. Un valor Z negativo indica que la observación está por debajo de la media; un valor Z positivo indica que la observación se encuentra por encima del valor de la media (Rosenshein et al., 2011). Los valores Z son simplemente desviaciones estándar. Si, por ejemplo, se obtiene un valor Z de +2,5, se entiende que el resultado son desviaciones estándar de 2,5. Por otro lado, el valor P sirve para medir el nivel de significación de la prueba. Si el nivel de significación es superior al valor p se debe rechazar la hipótesis nula.

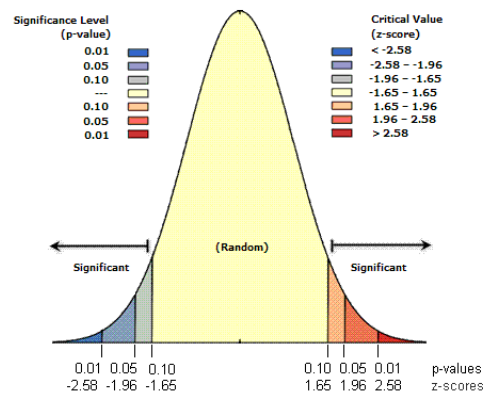
Para las herramientas de análisis de patrones, existe la probabilidad de que el patrón espacial observado se haya creado mediante algún proceso aleatorio. Cuando el valor p es muy pequeño, significa que es muy poco probable que el patrón espacial observado sea el resultado de procesos aleatorios, por lo tanto, puede rechazar la hipótesis nula. Para determinar esto se puede utilizar la tabla a continuación:

Tabla 5.13: Valor P y valor Z críticos para distintos niveles de confianza.

Valor Z (Desviaciones estándar)	Valor P (Probabilidad)	Nivel de confianza
< -1,65 o > +1,65	< 0,10	90%
< -1,96 o > +1,96	< 0,05	95%
< -2,58 o > +2,58	< 0,01	99%

Tanto los valores Z como el valor P se asocian con la distribución normal estándar. Un valor Z muy alto o muy bajo (negativo), asociados con un P valor muy pequeño, se encuentra en la cola de la distribución normal (Rosenshein et al., 2011). Esto es indicativo que es poco probable que el patrón espacial observado refleje el patrón aleatorio teórico representado por su hipótesis nula. Para rechazar la hipótesis nula, se debe formular una opinión subjetiva con respecto al grado de riesgo que desea aceptar por estar equivocado. Por lo tanto, antes de ejecutar la estadística espacial, se selecciona un nivel de confianza.

Gráfica 5.5: Distribución de valor Z y P Valor.



Fuente: <http://help.arcgis.com>

Al examinar los resultados obtenidos (Tabla 5.15) se observa que los valores Z para los cinco años analizados superan el +2,58. Todos los resultados muestran una tendencia a la concentración. Dado estos valores se debe entender que existe menos del 1% de probabilidad que los resultados sean productos del azar. Por otro lado, los resultados del Valor Z son todos positivos, indicativo que las observaciones están por encima de la media.

TABLA 5.14: Índice Moran

	2005	2006	2007	2008	2009
Índice Moran	0,160030	0,292665	0,374723	0,102181	0,079569
Valor Z	6,845843	15,765071	2,589689	4,601856	2,907656
Valor P	0,000	0,000	0,009606	0,000004	0,003641

Los valores residenciales tienden a estar autocorrelacionados por una serie de condiciones. Entre ellas se encuentra que los vecindarios se desarrollan al mismo tiempo lo que resulta en residencias que cuentan con características estructurales, de localizaciones y de vecindario similares. Además, los valores residenciales tienden a evolucionar de manera paralela a la de las residencias vecinas lo que hace que los precios de inmuebles no sean independientes a los de su vecindario.

Las pruebas diagnósticas realizadas muestran varios problemas, principalmente uno de alta correlación espacial en las ventas. Ya se planteó anteriormente que los residuales producidos por un modelo de valoración hedónica aplicado a inmuebles pueden estar espacialmente correlacionados por tres razones (Gillen et al., 2001; Tu et al., 2004; Suriatini, 2006). Primero por las características estructurales del inmueble y las características del vecindario donde se localiza, segundo por el proceso de valoración utilizado para determinar el precio del inmueble y tercero por malas especificaciones del modelo MCO.

Basado en estos resultados se puede asumir que existe un problema de predictores redundantes (Rosenshein et al., 2011). Esa redundancia puede dar paso a la determinación de autocorrelación espacial por las respectivas pruebas. En los datos utilizados para el análisis se registran ventas en urbanizaciones nuevas. Dependiendo del tipo de inmueble en un período de venta de un año se puede entregar entre 20 y 100 unidades de vivienda en un nuevo complejo residencial. En ocasiones esta cifra puede ser superior. Dada las características de los inmuebles, los casos utilizados para el análisis muestran poca o ninguna variación en sus características estructurales. Por otro lado, se observa una alta multicolinealidad en algunas variables del modelo. La multicolinealidad es una característica de la muestra seleccionada y no de los errores de estimación. Uno de los factores que impulsa a esta condición cuando alguna o todas las variables explicativas están altamente correlacionadas entre sí.

La utilización de la totalidad de ventas de inmuebles para cada año de estudio puede provocar estos resultados irregulares. Para corregir esta situación se decidió utilizar un método donde se trabajará con una muestra de los datos y no el universo. Para esto se realizaron dos muestras de los datos. La primera fue obtenida por medio de un muestreo aleatorio y el segundo una representación de los valores medianos de venta y características de cada comunidad. Con esto se busca eliminar la redundancia de casos que afectan los resultados esperados. La muestra aleatoria se logró con la herramienta de muestreo que contiene las herramientas de trabajo Hawth's Tools⁸. Ésta realiza una selección aleatoria de casos de una base de datos digital cuya representación espacial se da por medio de puntos. La cantidad de casos seleccionada corresponde a un porcentaje del total establecido por el usuario. En el caso de estudio se estableció que la muestra fuera un 30% del total de casos. La segunda muestra se constituyó de la selección de ventas que representaran el valor mediano para cada una de las comunidades identificadas en la zona de estudio. Dada la alta similitud en las

⁸ Se puede descargar en www.spatial ecology.com

características de los inmuebles en cada comunidad se asume que el inmueble cuyos valores se consideren como la mediana de todos los valores de venta es representativo de la comunidad.

Con estas dos selecciones de datos se procedió a realizar una comparación para determinar cuál mostraba el mejor desempeño. La comparación se fundamentó en tres factores. Primero el valor R^2 , segundo el nivel de autocorrelación espacial y tercero el Factor de Inflación de Varianza. Bajo las tres condiciones el mejor desempeño en los cinco años analizados se observó en los valores que representaban el valor mediano de las ventas por cada comunidad. Dado estos resultados se procedió a realizar el análisis exploratorio de datos, esta vez con la muestra seleccionada.

5.3 Diagnóstico del modelo de regresión con datos muestreados (Valor Mediano)

La primera prueba realizada fue la del Coeficiente de determinación múltiple R^2 . Como se indicó anteriormente esta prueba indica cuánta variación de la variable dependiente es explicada por el modelo. Los modelos para el período bajo estudio muestran valores R^2 que fluctúan entre 0,935 y 0,966 (Tabla 5.15). En otras palabras, en este período de tiempo las variables independientes seleccionadas pueden explicar entre un 93% y un 96% el comportamiento de la variable dependiente. Esto es cónsono con lo establecido en otros estudios (Orford, 1999) y mejora lo obtenido en otros (Yu, 2007).

Al comparar con los valores R^2 obtenidos con el universo de datos se puede observar que superan los mismos. Se toma como acertada la selección de las variables explicativas escogidas ya que pueden explicar 93% o más del valor de las propiedades utilizadas en el estudio.

Tabla 5.15: Coeficiente de Determinación Múltiple (R^2): Universo vs Valor Mediano.

Prueba	2005	2006	2007	2008	2009
Coeficiente R^2 (Universo)	0,953	0,924	0,937	0,929	0,898
Coeficiente R^2 (Valor Mediano)	0,966	0,945	0,955	0,935	0,940

Como segundo paso en el proceso de diagnóstico de los datos muestreados se evaluaron las variables explicativas utilizadas en el modelo analizando el resultado del coeficiente obtenido (Tabla 5.16). Cada variable explicativa refleja la fortaleza y el tipo de relación que tienen las variables explicativas con la variable dependiente. Para analizar este aspecto se construyó una tabla (Tabla 5.17) indicando si el valor esperado para cada variable explicativa era positivo o negativo y se comparó con el resultado obtenido para los 5 años de estudio. Se esperaba que 16 variables obtuvieran un coeficiente positivo y tres un coeficiente negativo. En ninguno de los 5 años analizados se observó el patrón esperado. Sí se pudo observar variables particulares que se comportaron al largo de los cinco años según lo esperado. En términos de variables explicativas de tipo positivo solo la superficie del solar (LOTAREA), la superficie de ocupación (HouseareaM), la piscina (SWP), la cancha de Voleibol (VB), el parque de béisbol (BB) y el campo de golf (GLF) arrojaron coeficientes según esperados en los cinco años. En comparación con los resultados obtenidos con el modelo que utiliza el universo de las ventas, solo la superficie del solar (LOTAREA), la superficie de ocupación (HOUSEAREAM), la cancha de voleibol (VB) y el parque de béisbol (BB) mostraron coeficientes con un comportamiento similar.

Por su parte, de las variables explicativas con efecto negativo, la edad del inmueble (AGE) arrojó un coeficiente esperado en los cinco años. En el diagnóstico utilizando el universo de las ventas esta variable presentó el mismo comportamiento.

Las restantes variables mostraron coeficientes que variaban entre positivos y negativos en los diversos años. Las que más resaltan son la de cantidad de baños (BATHROOMS) y gimnasio (GYM). En el caso de cuartos de baño se mantiene el efecto de confusión en la interpretación del signo (Kennedy, 2005). En tres de los cinco años analizados el coeficiente fue negativo. De igual manera, la existencia de facilidades de gimnasio en las comunidades siempre es considerada un punto a favor de la venta del inmueble y que encarece el valor de la propiedad; no obstante, el coeficiente fue negativo en cuatro de los cinco años para esta variable.

Es importante señalar que el año 2007 muestra los resultados de coeficientes más similares a los esperados. Sin embargo, a partir de ese mismo año se observa que este patrón varía significativamente de un año a otro. Se puede asumir que la ruptura de la burbuja inmobiliaria en ese momento hizo que los patrones de venta y consumo de inmuebles variaran mucho entre los consumidores de inmuebles. La disposición a pagar por lujos en las residencias pasa

a un segundo plano ante la incertidumbre que se crea en el mercado inmobiliario. Por otro lado, la disposición a vender inmuebles a precios muy por debajo de lo que había establecido el mercado se ve reflejado en estos resultados.

Tabla 5.16: Coeficientes (Valor Mediano)

Variable	2005	2006	2007	2008	2009
LOTAREA	53,372	9,883	161,552	26,949	117,663
AGE	-622,213	-928,356	-1074,563	-392,241	-2339,752
BEDROOMS	-10105,311	-15698,055	-5412,943	-8438,369	6821,527
BATHROOMS	-3704,522	-348,397	7591,833	-547,729	19970,035
CC_DIST	6,401	-,595	-1,648	-7,243	-10,319
EXIT_DIST	-5,658	5,983	2,171	11,487	-2,507
HOUSEAREAM	1102,966	1685,951	596,976	1336,248	306,138
ACC	32164,448	8856,299	14063,552	22680,655	-27847,705
GZB	14871,193	-12085,571	18792,046	7171,259	26921,602
PG	-8237,690	4332,981	3578,716	-3825,442	-10025,387
SWP	4775,205	24160,807	3060,267	18940,130	5533,707
TENIS	30424,177	26575,136	36906,005	-3303,205	26744,583
BSKT	5789,556	-19307,323	-12500,671	-2543,280	-12737,214
PP	-4603,701	894,626	6022,168	-15517,067	-24101,598
GYM	-44113,540	-35534,943	30793,986	-10234,425	-42391,260
CH	26464,823	13563,394	6803,431	-6815,979	48556,474
VB	53335,635	19933,875	14173,144	45346,408	56532,337
BB	19187,542	32936,945	8425,752	18437,162	12455,415
GLF	334672,282	285112,362	127333,996	95438,549	37835,230

Tabla 5.17: Relaciones Esperadas y obtenidas por año (Positivas / Negativas) (Valor Mediano).

Variable	Esperada	2005	2006	2007	2008	2009
LOTAREA	P	P	P	P	P	P
AGE	N	N	N	N	N	N
BEDROOMS	P	N	N	N	N	P
BATHROOMS	P	N	N	P	N	P
CC_DIST	N	P	N	N	N	N
EXIT_DIST	N	N	P	P	P	N
HOUSEAREAM	P	P	P	P	P	P
ACC	P	P	P	P	P	N
GZB	P	P	N	P	N	P
PG	P	N	P	P	N	N
SWP	P	P	P	P	P	P
TENIS	P	P	P	P	N	P
BSKT	P	P	N	N	N	N
PP	P	N	P	P	N	N
GYM	P	N	N	P	N	N
CH	P	P	P	P	N	P
VB	P	P	P	P	P	P
BB	P	P	P	P	P	P
GLF	P	P	P	P	P	P

Una vez analizado el comportamiento de los coeficientes se procedió a verificar la Probabilidad y la Probabilidad Robusta. Estas miden la significancia estadística del coeficiente. En el informe de resultados generado se coloca un asterisco a las probabilidades estadísticamente significativas. Una variable explicativa asociada a un coeficiente estadísticamente significativo es importante para el modelo de regresión, si el sentido común asume una relación válida con la variable dependiente.

Al observar los resultados para los cinco años bajo estudio se puede observar que la única variable significativa en los cinco años analizados fue existencia de control de acceso para la comunidad (ACC) (Tabla 5.18). Con tres años o más mostrando variables significativas se encuentra la superficie del solar medido en metros cuadrados (LOTAREA) y la cantidad de dormitorios (BEDROOMS). En cuanto a las variables con probabilidades mayores a 0,05% fueron gazebo (GZB), área de juego para niños (PG), la piscina (SWP), cancha de tenis (Tenis), cancha de baloncesto (BSKT), parque pasivo (PP), gimnasio (GYM) y cancha de voleibol (VB). Las restantes variables se pueden dividir entre aquellas que solo presentaron

valores significativos en dos años o menos. Entre éstas se puede mencionar la edad del inmueble (AGE) o superficie del solar (LOTAREA).

Tabla 5.18: Variables Significativas, Probabilidad menor a 0,05 (95% de confianza) (Valor Mediano)

Variable	2005	2006	2007	2008	2009
LOTAREA	0,014736*	0,049914*	0,049575*	0,877324	0,611875
AGE	0,476953	0,393213	0,000000*	0,000055*	0,543063
BEDROOMS	0,069444	0,153137	0,001054*	0,017343*	0,008686*
BATHROOMS	0,009322*	0,000738*	0,501982	0,719092	0,091168
CC_DIST	0,302956	0,983521	0,4979	0,914482	0,026363*
EXIT_DIST	0,217902	0,671158	0,174204	0,016137*	0,083397
HOUSEAREAM	0,860776	0,919174	0,335048	0,046770*	0,175568
ACC	0,000000*	0,000000*	0,000345*	0,000000*	0,000196*
GZB	0,095358	0,548729	0,70143	0,659958	0,64496
PG	0,259052	0,5476	0,346484	0,891646	0,375908
SWP	0,321279	0,997153	0,874006	0,149885	0,101052
TENIS	0,657147	0,155868	0,85374	0,674751	0,446452
BSKT	0,05907	0,065157	0,265637	0,702503	0,627322
PP	0,695055	0,283237	0,276867	0,739824	0,789886
GYM	0,616484	0,972863	0,860773	0,644025	0,181392
CH	0,007192*	0,193508	0,719807	0,129209	0,571749
VB	0,112514	0,384069	0,981034	0,475182	0,855471
BB	0,005822*	0,252389	0,461724	0,070878	0,415642
GLF	0,043497*	0,003797*	0,750414	0,05537	0,44302

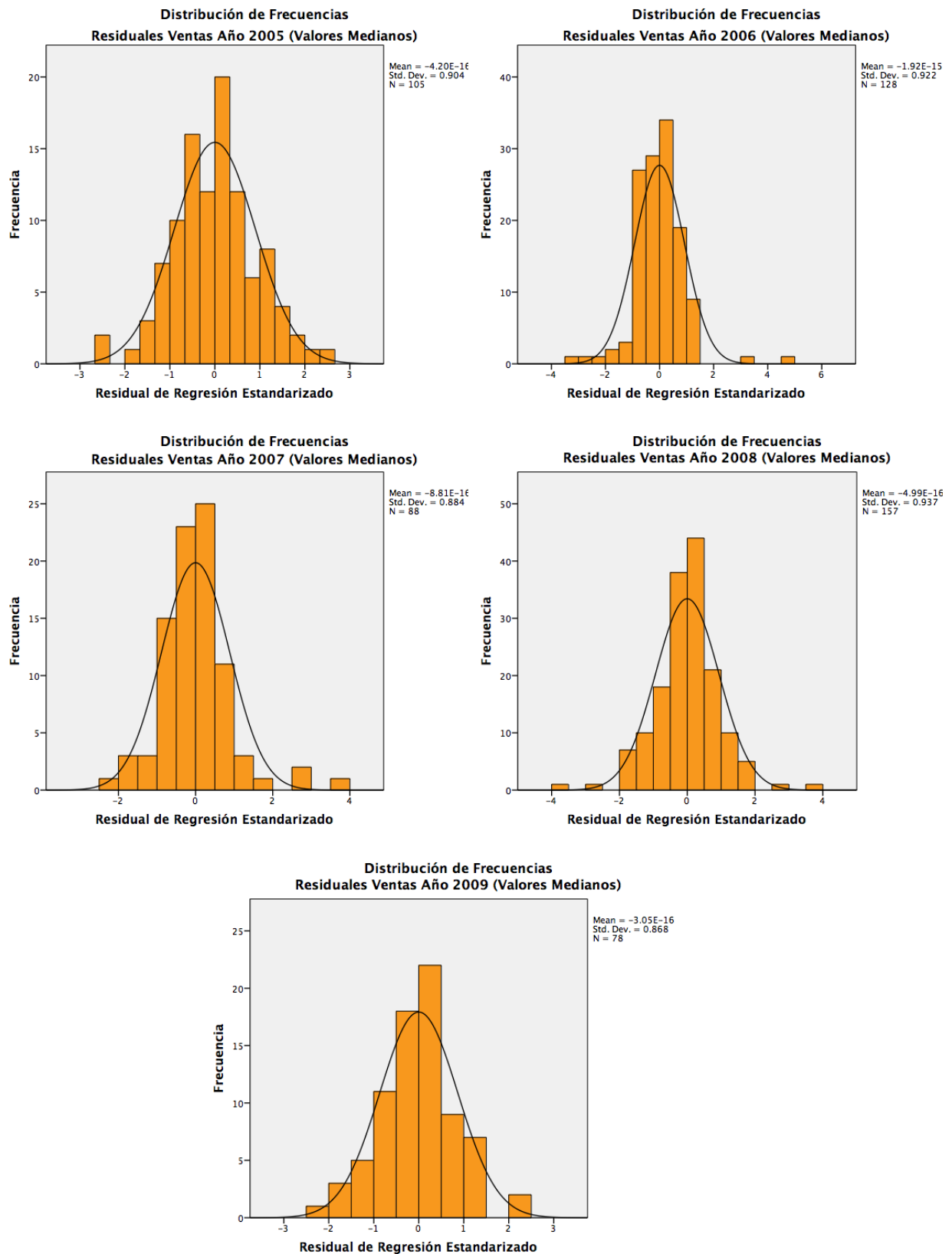
En cuanto a la multicolinealidad del modelo, al observar los valores del Factor de Inflación de la Varianza (FIV) (Tabla 5.19) solo se pudo identificar un caso que se reflejara esta condición. Se trata de la superficie de ocupación en metros (HOUSEAREAM) para el año 2005. El valor del FIV para esta variable fue de 9,210, superior al 7,5 que se establece como límite superior aceptable. En los restantes años se pudo identificar valores FIV altos, pero ninguno superaba el límite superior adoptado.

En términos de la distribución de los valores residuales, al analizar el coeficiente de distribución se entiende que la distribución es normal (Tabla 5.20). Las simetrías son todas positivas con excepción a la del año 2008 que es simétrica. Estableciendo el límite del $\pm 0,5$, tan solo los coeficientes de asimetría para el 2005 y 2008 se localizarían dentro de este rango. En cuanto a la curtosis todas las distribuciones serían leptocúrticas ya que superan el valor de cero de manera positiva.

TABLA 5.19: VIF – Factor de Inflación de Varianza de Valores Medianos (7,5)

Variable	2005	2006	2007	2008	2009
LOTAREA	2,879	1,813	2,618	2,196	3,504
AGE	3,150	3,123	4,113	3,187	3,900
BEDROOMS	2,817	1,865	2,254	1,631	2,310
BATHROOMS	4,263	4,248	3,947	3,275	5,164
CC_DIST	3,536	3,342	3,766	3,215	3,528
EXIT_DIST	4,749	3,863	4,467	3,414	3,533
HOUSEAREAM	9,210	5,422	5,709	6,529	7,447
ACC	4,378	4,460	4,531	5,065	6,852
GZB	4,719	4,617	5,516	7,624	6,778
PG	4,254	4,233	4,253	3,588	5,390
SWP	3,887	3,694	5,745	8,800	5,058
TENIS	2,788	2,644	3,262	2,572	2,727
BSKT	1,935	1,943	1,886	2,002	2,346
PP	1,920	1,942	2,082	2,408	1,755
GYM	6,157	4,629	7,225	4,911	6,870
CH	4,214	3,601	5,073	5,194	6,580
VB	2,645	2,649	2,676	2,262	2,215
BB	1,600	1,458	1,596	1,584	1,634
GLF	1,407	1,184	1,885	1,642	1,702

Gráfica 5.6: Histogramas de Residuales, 2005 – 2009 (Valor Mediano por Comunidad)



Los datos muestran una distribución normal (Gráfica 5.6). En cuanto a la curtosis, todas las distribuciones son leptocúrticas ya que superan el valor de cero de manera positiva.

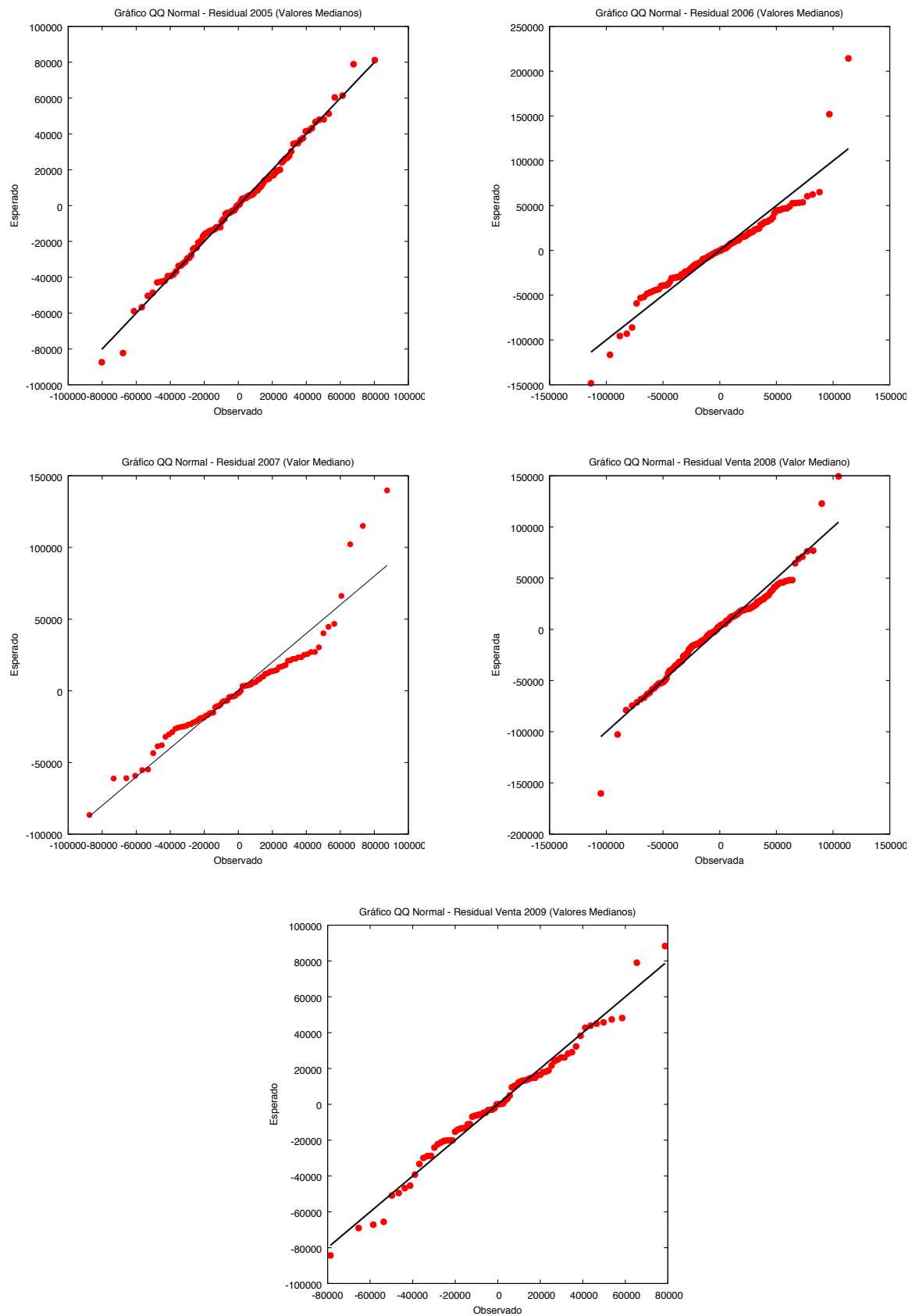
Al revisar los resultados de la prueba Jarque – Bera (JB) (Tabla 5.20) se pudo observar que en los años 2005 y 2009 se dan valores por debajo de los 5,99 utilizados como regla de decisión. Además, se observa que los valores p para estos años no es estadísticamente significativa. Esto significa que solo para el año 2005 y el 2009 no se rechaza la hipótesis nula. Se puede concluir que los residuos del modelo de regresión no están distribuidos normalmente con excepción de los años indicados.

TABLA 5.20: Resultado Jarque – Bera (Mediana)

	2005	2006	2007	2008	2009
Jarque-Bera (Universo)	396,108841	2049,711556	211,624527	194,613170	41,293718
Valor P (Universo)	0,000000*	0,000000*	0,000000*	0,000000*	0,000000*
Jarque-Bera (Mediano)	0,638	251,139	70,000	10,556	1,335
Valor P (Mediano)	0,726	0,000000*	0,000000*	0,000510*	0,5127

Por otro lado, al interpretar la distribución del conjunto de datos y su similitud a la distribución ideal por medio de los gráficos Cuantil-Cuantil, se puede observar que la del año 2005 presenta la distribución más similar a la ideal. En cada uno de los años restantes se muestran un punto muy alejado de la distribución en la parte superior de la gráfica, lo que lleva a pensar en la existencia de un valor extremo. Se puede observar un poco de movimiento alrededor de la línea, sin embargo, esto no debe descalificar a estos datos de estar distribuidos normalmente.

Gráfica 5.7: Gráficos QQ Normal – Residuales Ventas 2005-09 (Valores Medianos)



La estacionariedad del modelo fue evaluada con el índice estadístico de Koenker (Índice estadístico de Breusch-Pagan estudentizado de Koenker - BP). Como se indicó anteriormente, ésta es una prueba utilizada para determinar si las variables explicativas del modelo tienen una relación consistente con la variable dependiente, tanto en el espacio geográfico como en el espacio de datos. La hipótesis nula de esta prueba es que el modelo es estacionario. Al calcular la prueba si el valor de Koenker es bajo y el P Valor significativo, las relaciones que se están modelando presentan una no-estacionariedad estadísticamente significativa. En otras palabras, las relaciones entre las variables varían a lo largo de la zona de estudio. Al observar los resultados de esta prueba en los cinco años analizados (Tabla 5.21) todos los p valores son menores al 0,05. Se asume que los modelos de cada año cuentan con una heterocedasticidad significativa y son no-estacionarios.

TABLA 5.21: Estacionariedad del modelo (Koenker BP) (Valor Mediano)

Prueba	2005	2006	2007	2008	2009
Koenker – (Breush – Pagan)	38,363	39,51	31,330	45,474	40,606
Valor P	0,0053*	0,00378*	0,03712*	0,00058*	0,00272*

Dentro del análisis del desempeño del modelo, la determinación de autocorrelación espacial se considera uno de los aspectos más importantes. Para evaluar la autocorrelación espacial residual se calculó el Índice de Moran. Éste mide la autocorrelación espacial basada en la localización de los elementos y su respectivo valor de manera simultánea.

Gráfica 5.8: Resultados del índice de Moran Con Valores Medianos

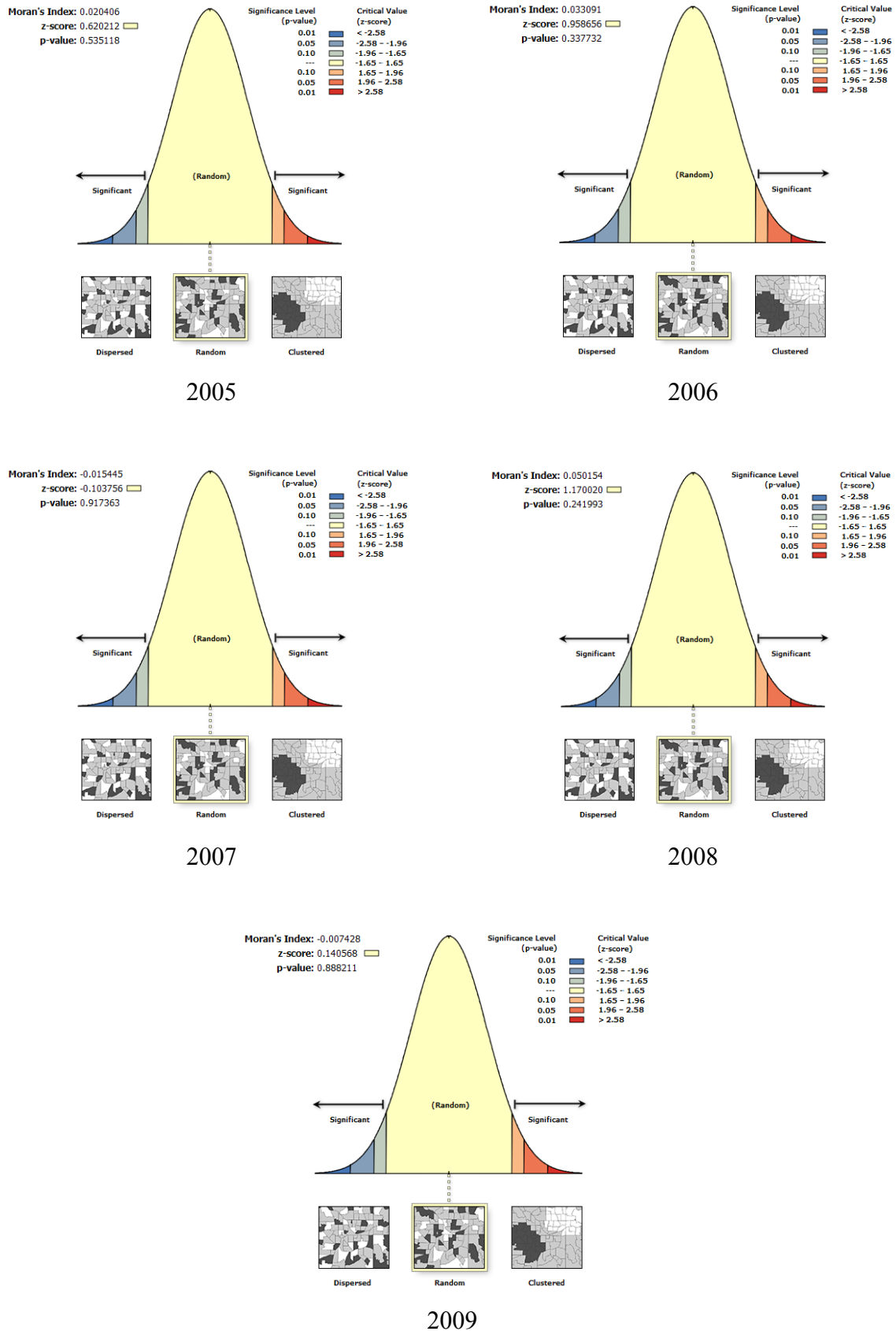


TABLA 5.22: Índice Moran (Valor Mediano)

	2005	2006	2007	2008	2009
Índice Moran	0,020406	0,033091	-0,015445	0,050154	-0,007428
Valor Z	0,620212	0,958656	-0,103756	1,170020	0,140568
P - valor	0,535118	0,337732	0,917363	0,241993	0,888211

Al examinar los resultados obtenidos (Tabla 5.22) se observa que los valores del índice giran en torno al cero. Por ello, se entiende que los datos presentan un patrón espacial totalmente aleatorio. En cuanto al valor Z, solo el año 2007 muestra un valor negativo. Esto indica que la observación está por debajo de la media. Los años restantes muestran valores positivos, entendiéndose que las observaciones se encuentran por encima del valor de la media. Sin embargo, todos los valores observados se encuentran en el rango de -1,65 a 1,65 (ver Tabla 5.14) lo que establece que estos son productos de una distribución aleatoria y no existe autocorrelación espacial.

Como prueba final se procedió a comparar los modelos utilizando el Criterio de información de Akaike corregido (AICc). Como se puede ver el proceso de construcción del modelo es iterativo y somete a prueba modelos diferentes. El modelo con el valor del AICc más pequeño es el mejor, es decir, teniendo en cuenta la complejidad del modelo el modelo con el AICc más pequeño se ajusta mejor a los datos observados.

TABLA 5.24: Prueba AiCc

Datos	2005	2006	2007	2008	2009
Originales	12371,61	21313,30	5490,27	11166,74	5538,75
Valor Mediano	2508,40	3133,53	2128,01	3799,15	1876,67

Los valores del Criterio de información de Akaike obtenidos para los modelos basados en el valor mediano (Tabla 5.24) arrojaron valores menores a los obtenidos para los modelos donde se utilizó el universo de las ventas por año. En el caso de este último, los valores del criterio más bajos fueron para los años 2007 y 2009 (5502,412056 y 5578,348302, respectivamente).

Sin embargo, los valores para estos años utilizando los valores medianos fueron mucho más bajos (2227,648640 y 2023,781558, respectivamente).

Los resultados obtenidos de la ejecución de estas pruebas diagnósticas utilizando los datos de valor mediano para cada comunidad, complejo residencial y sector de la zona de estudio lleva a concluir que el desempeño general de los modelos es superior a que utiliza el universo de los datos. Queda de esa manera descartado el uso del universo de los datos para el modelo y se decide seguir realizando pruebas buscando un mejor desempeño del modelo.

5.4 Selección Final de Variables

El proceso de análisis exploratorio evidenció que la creación de modelos con las variables seleccionadas mostraba un desempeño mejor cuando se utiliza el valor mediano de ventas residenciales por cada comunidad versus la selección del universo de ventas reflejada en cada comunidad. A pesar del marcado desempeño entre el modelo con el universo de venta versus el valor mediano, el comportamiento individual de las variables en cada modelo es muy variado. Los resultados del coeficiente de determinación múltiple R^2 indican que las variables independientes pueden explicar un mayor porcentaje del valor del inmueble cuando se utilizan los datos del valor mediano por comunidad. Por otro lado, la prueba del Criterio de Información de Akaike (AICc) confirma que el modelo con las ventas medianas se ajusta mejor a los datos observados. Con los resultados de la prueba del Índice de Moran no se rechaza la hipótesis nula que establece que el atributo que se analiza está distribuido en forma aleatoria entre las entidades del área de estudio.

Los resultados de estas pruebas son indicativos de una selección correcta de datos. Sin embargo, los resultados de otras pruebas muestran problemas con estos. Uno de estos aspectos es el de la multicolinealidad. Los datos muestran que los años 2006 al 2009 están libres de multicolinealidad, no así el año 2005. Una variable tan importante como la superficie de ocupación del inmueble arroja un valor de 9.210 en la prueba del Factor de Inflación de la Varianza (FIV). Esta cifra supera el 7,5 que se adopta como límite superior aceptable para determinar si la variable muestra multicolinealidad o no. Otro problema es el de la heterocedasticidad. Los valores del Índice estadístico de Breusch-Pagan estudentizado de Koenker - BP señalan una no estacionariedad significativa. Finalmente, solo se puede asumir la existencia de una distribución normal de los datos para algunos de los años bajo

análisis. Al ejecutar la prueba Jarque-Bera se observa que solo los años 2005 y 2009 cumplen con los parámetros de una distribución normal.

Una de las condiciones que puede incidir en este comportamiento es el problema de selección de las variables independientes. Con el objetivo de depurar a un grado mayor e identificar la combinación idónea de variables a utilizar en el modelo de predicción se aplicó a los datos el proceso de “Regresión Exploratoria” que contiene el programa ArcGIS. Esta herramienta evalúa todas las posibles combinaciones de predicción para una variable dependiente utilizando las variables independientes seleccionadas. Para ejecutar esta función se establecen los valores aceptados para cuatro pruebas exploratorias de datos que se realizan utilizando diversas combinaciones de variables independientes. Se busca que el modelo seleccionado cumpla con las siguientes condiciones:

- Que exceda un límite de R^2 Ajustado especificado
- Que los valores P de los coeficientes sean menores a los especificados
- Que los valores de coeficiente de FIV, para todas las variables explicativas, sea menor que el especificado
- Que cuente con un valor P Jarque-Bera más grande que el especificado

Para ejecutar la herramienta se seleccionaron las 19 variables bajo análisis y se indicó que los parámetros a cumplir serían los siguientes:

- R^2 Ajustado: Mayor a 0,70
- FIV: Menor a 7,5
- P-valor: máximo de 0,05
- Valor P Jarque Bera: Mínimo aceptado 0,1
- Valor P Autocorrelación espacial: Mínimo aceptado 0,1

Una vez ejecutado el análisis exploratorio, se observó la selección de nueve variables de manera consistente en modelos para los diversos años. Tomando estas nueve variables se ejecutó nuevamente la técnica de los Mínimos Cuadrados Ordinarios para cada año, arrojando valores aceptables para los cinco años de estudio (Tabla 5.25). Se estableció así que éstas serían las variables finales para la comparación de técnicas. Las variables son:

- Estructurales de la residencia
 - LOTAREA: Superficie del solar medido en metros cuadrados.
 - AGE: Años que lleva construido el inmueble.
 - HOUSEAREAM: Superficie de ocupación del inmueble medido en metros cuadrados.
- Comunitarios
 - TENIS: Existencia de canchas de tenis
 - GYM: Existencia de facilidades de gimnasio para ejercicios
 - VB: Existencia de cancha de voleibol.
 - BB: Existencia de parque de béisbol.
 - GLF: Existencia de campo de golf.
- Distancia a servicios (Localización)
 - CC_Dist: Distancia euclidiana a centro comercial más cercano

TABLA 5.25: Resultados Modelo de 9 Variables Finales

	2005	2006	2007	2008	2009
R2	0,91	0,88	0,90	0,85	0,86
VIF (Mas Alto)	2,86	2,53	2,29	2,56	3,22
AiCc	2517,24	3131,70	2122,16	3799,74	1875,24
Koenker (P valor)	0,0540	0,000047*	0,000934*	0,0000*	0,00005*
Jarque Bera (P Valor)	0,8134	0,0000*	0,0000*	0,0000*	0,08043
Moran I	0,0711	-0,3292	-0,0224	0,0597	0,0408
Valor Z	1,039139	-0,546229	-0,15598	0,741716	0,764689
Valor P	0,29874	0,58490	0,87604	0,45825	0,44457

Con la selección de las nueve variables se cuenta con un modelo propiamente establecido. Se procede a realizar la Regresión Geográficamente Ponderada para cada año y comparar los resultados con los obtenidos de la prueba de los Mínimos Cuadrados Ordinarios.

Capítulo 6

Valoración Hedónica con Mínimos Cuadrados Ordinarios y Regresión Geográficamente Ponderada: Resultados de un Estudio Comparado

6. Valoración Hedónica con Mínimos Cuadrados Ordinarios y Regresión Geográficamente Ponderada: Resultados de un Estudio Comparado

La comparación de los modelos de Mínimos Cuadrados Ordinarios y Regresión Geográficamente Ponderada busca adoptar un esquema que permita la actualización de los valores inmuebles que utilizan las agencias gubernamentales de Puerto Rico basado en información de ventas inmuebles que se producen cada año. La comparación se realiza para identificar cuál de los dos modelos permite producir una predicción más acertada. Ya seleccionadas las nueve variables a emplear en el modelo se procedió a ejecutar los mismos y aplicar una serie de pruebas diagnósticas que permitirían la comparación y la eventual selección del modelo en función de su desempeño.

6.1 Construcción del modelo con la técnica de los Mínimos Cuadrados Ordinarios

Tomando las pruebas diagnósticas discutidas en el capítulo anterior se procedió a construir el modelo utilizando la técnica de los mínimos cuadrados. Se parte el análisis con la prueba del Coeficiente de determinación múltiple R^2 . Para el período bajo estudio se observan valores R^2 que fluctúan entre 0,85 y 0,91 (Tabla 6.1). En otras palabras, en este período de tiempo las variables independientes seleccionadas pueden explicar entre un 85% y un 91% el comportamiento de la variable dependiente.

Tabla 6.1: Coeficiente de Determinación Múltiple (R^2)

Prueba	2005	2006	2007	2008	2009
Coeficiente determinación múltiple R^2	0,91	0,88	0,90	0,85	0,86

Como segundo paso en el diagnóstico del modelo se evaluaron las variables explicativas utilizadas en este analizando el resultado del coeficiente obtenido (Tabla 6.2). Cada variable explicativa refleja la fortaleza y el tipo de relación que tienen las variables explicativas con la

variable dependiente. Para analizar este aspecto se construyó una tabla (6.3) indicando si el valor esperado para cada variable explicativa era positivo o negativo y se comparó con el resultado obtenido para los 5 años de estudio. Se esperaba que siete variables obtuvieran un coeficiente positivo y dos un coeficiente negativo. Solo en los años 2007 y 2009 se observó el patrón esperado. En los restantes años todas las variables presentaron resultados de acuerdo a lo esperado con la excepción de la variable gimnasio (GYM). Se esperaba que los resultados de esta fueran positivos, presentando esa condición solo en los años 2007 y 2009.

Por su parte, de las variables explicativas con efecto negativo solo la edad del inmueble (AGE) arrojó un coeficiente según esperado en los cinco años. La otra variable que se esperaba arrojara un coeficiente negativo era distancia a centro comercial más cercano (CC-Dist). Este fue negativo en todos los años menos en el 2006.

Tabla 6.2: Coeficientes

Variable	2005	2006	2007	2008	2009
LOTAREA	50,5382	7,16378	156,253	30,825	109,866
AGE	-1742,21	-1364,17	-1762,12	-1170,69	-2027,47
CC_DIST	-1,16594	3,69053	-1,80957	-1,64349	-12,9612
HOUSEAREAM	1034,02	1595,39	669,52	1253,89	545,186
TENIS	41231,3	30573,1	45699,8	13954,8	31473,6
GYM	-28780,1	-10779	45236	-240,186	1751,49
VB	54888,8	9280,31	10409,8	48038,2	43352,7
BB	11046,3	31129,3	5003,67	13154	7568,15
GLF	326179	283477	114167	98155,6	16193,7

Tabla 6.3: Relaciones Esperadas y obtenidas por año (Positivas / Negativas)

Variable	Esperada	2005	2006	2007	2008	2009
LOTAREA	P	P	P	P	P	P
AGE	N	N	N	N	N	N
CC_DIST	N	N	P	N	N	N
HOUSEAREAM	P	P	P	P	P	P
TENIS	P	P	P	P	P	P
GYM	P	N	N	P	N	P
VB	P	P	P	P	P	P
BB	P	P	P	P	P	P
GLF	P	P	P	P	P	P

A diferencia con los modelos exploratorios creados con mayor cantidad de variables explicativas, los resultados de coeficientes son similares a los esperados. Las nueve variables seleccionadas presentan un mejor comportamiento en cuanto a coeficientes.

Los valores del Factor de Inflación de la Varianza (FIV) (Tabla 6.4) indican que el modelo está ausente de multicolinealidad. Todos los valores se encuentran por debajo del 7,5 establecido como límite superior aceptable. El valor FIV más alto observado lo arrojó la variable superficie de ocupación en metros (HOUSEAREAM) para el año 2009. Este valor fue de 3,229, valor menor al 7,5 que se establece como límite superior aceptable. El valor FIV más bajo fue 1,111 para el campo de golf en el año 2006.

TABLA 6.4: VIF – Factor de Inflación de Varianza

Variable	2005	2006	2007	2008	2009
LOTAREA	2,141	1,647	2,294	1,922	2,317
AGE	1,660	1,683	2,073	1,859	2,185
CC_DIST	1,382	1,176	1,400	1,561	1,504
HOUSEAREAM	2,869	2,534	2,195	2,569	3,229
TENIS	2,111	2,022	2,122	1,988	1,853
GYM	2,528	2,287	2,649	2,189	1,887
VB	2,139	2,156	1,852	1,705	1,545
BB	1,425	1,288	1,425	1,409	1,211
GLF	1,276	1,111	1,571	1,356	1,386

Al observar los gráficos de distribución de frecuencia de los residuales (Gráfica 6.1) se puede observar que los años 2005 y 2009 muestran una distribución normal de los datos.

Esto se puede evidenciar con la prueba Jarque-Bera. Al revisar los resultados de esta prueba (Tabla 6.5) se puede observar que los valores p para los años 2005 y 2009 no son estadísticamente significativos. Esto significa que solo para estos años no se rechaza la hipótesis nula. Se puede concluir que los residuos del modelo de regresión no están distribuidos normalmente con excepción de los años indicados.

Gráfica 6.1: Histogramas de Residuales, 2005 – 2009 (Valor Mediano por Comunidad)

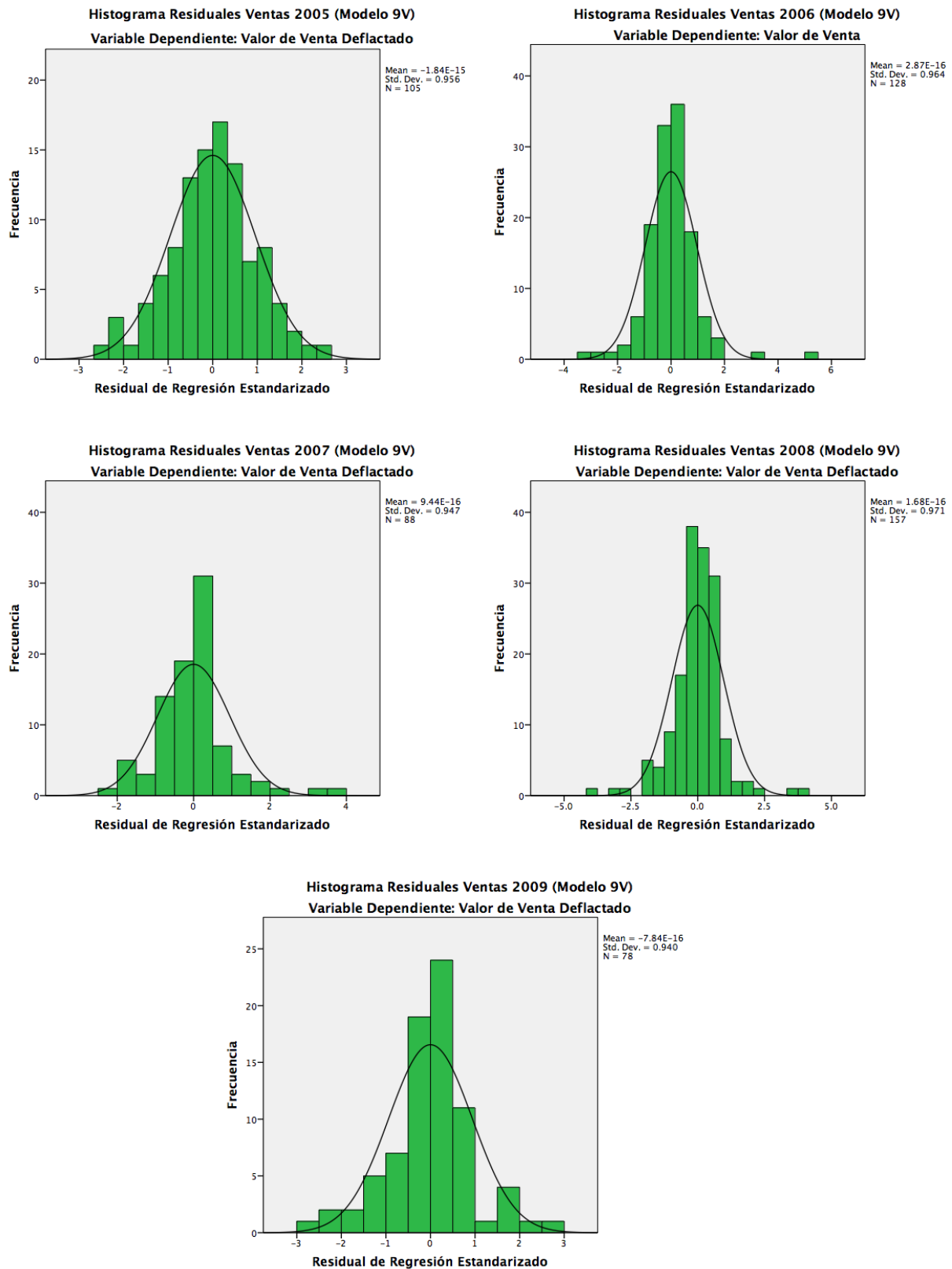


TABLA 6.5: Resultado Prueba Jarque – Bera

	2005	2006	2007	2008	2009
Jarque-Bera					
Valor P	0,8134	0,0000*	0,0000*	0,0000*	0,08043

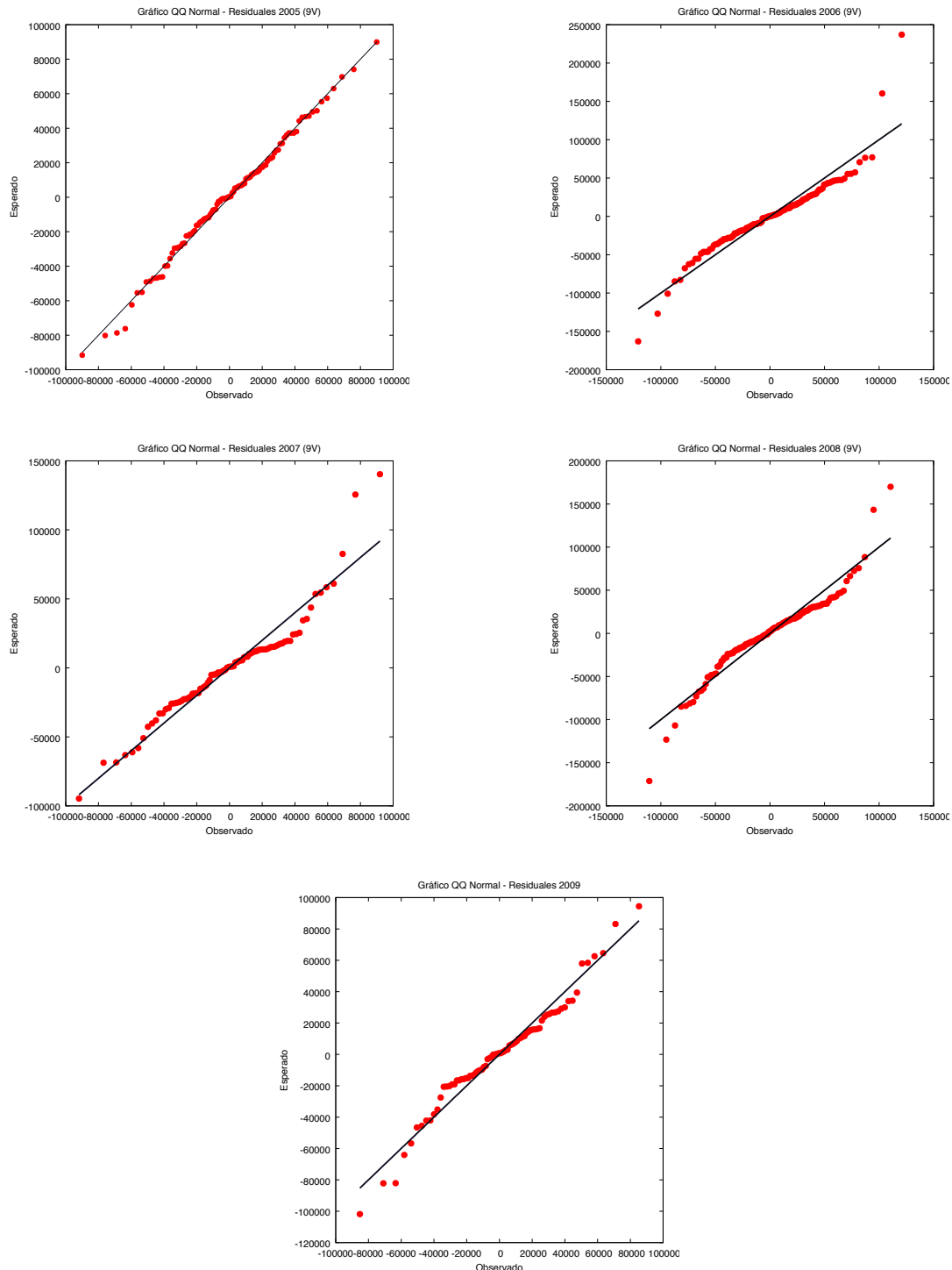
Por otro lado, al interpretar la distribución del conjunto de datos y su similitud a la distribución ideal por medio de los gráficos Cuantil-Cuantil (Gráfico 6.2), se puede observar que la del año 2005 presenta la distribución más similar a la ideal. En cada uno de los años restantes se muestran algunos puntos alejados de la distribución, principalmente en la parte superior de la gráfica. Esto lleva a pensar en la existencia de valores extremos. Se puede observar un poco de movimiento alrededor de la línea, sin embargo, esto no debe descalificar la distribución. Se puede asumir que es bastante similar a la distribución ideal esperada.

En cuanto a la estacionariedad del modelo, esta fue evaluada con el índice estadístico de Koenker (Índice estadístico de Breusch-Pagan estudentizado de Koenker - BP). Como se indicó anteriormente esta es una prueba utilizada para determinar si las variables explicativas del modelo tienen una relación consistente con la variable dependiente, tanto en el espacio geográfico como en el espacio de datos. La hipótesis nula de esta prueba es que el modelo es estacionario. Al observar los resultados de esta prueba en los cinco años analizados (Tabla 6.6) se observan resultados estadísticamente significativos para los años 2006 al 2009. Esto es indicativo que las relaciones de las variables son no estacionarias para estos años. Este es un patrón que a pesar de los cambios en la selección de variables se ha mantenido consistente.

TABLA 6.6: Estacionariedad del modelo (Koenker BP)

Prueba	2005	2006	2007	2008	2009
LM Koenker – (Breush – Pagan)	16,677929	35,575355	28,053308	48,152433	40,789482
Valor P	0,054000	0,000047*	0,000934*	0,000000*	0,000005*

Gráfica 6.2 Gráficos QQ Normal – Residuales Ventas 2005-09



Al igual que el modelo que toma como muestra la media con 19 variables explicativas, los gráficos Cuantil-Cuantil muestra que la distribución del año 2005 presenta la distribución más similar a la ideal. Los años 2006, 2007 y 2008 muestran valores extremos que afectan el patrón de relación expresado en la gráfica. El año 2009, con poca variación, muestra también

valores extremos moderados en comparación a los tres años que le preceden. Esto no debe descalificar a estos datos de estar distribuidos normalmente.

Al evaluar la autocorrelación espacial dentro del desempeño del modelo, el valor obtenido de la prueba del Índice de Moran (Tabla 6.7) muestra que los valores del índice giran en torno al cero. Por este resultado se entiende que los datos presentan un patrón espacial totalmente aleatorio. En cuanto al valor Z, solo los años 2006 y 2007 muestra valores negativos. Esto indica que la observación está por debajo de la media. Los años restantes muestran valores positivos entendiéndose que las observaciones se encuentran por encima del valor de la media. Sin embargo, todos los valores observados se encuentran en el rango de -1,65 a 1,65 lo que establece que estos son productos de una distribución aleatoria y no existe autocorrelación espacial.

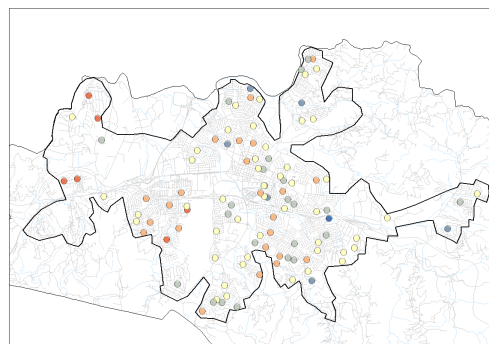
TABLA 6.7: Índice Moran

	2005	2006	2007	2008	2009
Índice Moran	0,0711	-0,3292	-0,0224	0,0597	0,0408
Valor Z	1,039139	-0,546229	-0,15598	0,741716	0,764689
P - valor	0,29874	0,58490	0,87604	0,45825	0,44457

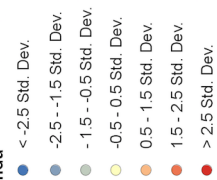
Estos resultados se pueden corroborar observando la distribución espacial de los residuales de la regresión. Cuando se cuenta con un modelo de regresión bien especificado se muestra una distribución aleatoria de sus residuales. Para los cinco años de estudio se puede observar esta condición (Gráfica 6.3).

Gráfica 6.3: Residuales de Regresión, Mínimos Cuadrados Ordinarios para los años 2005 al 2009

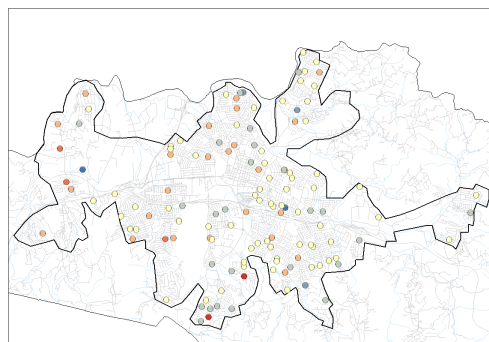
2005



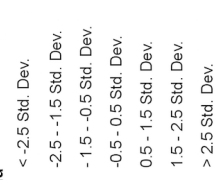
Leyenda



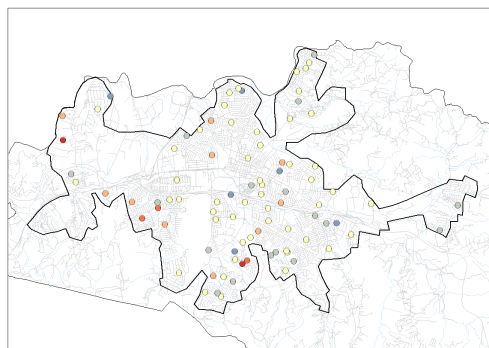
2006



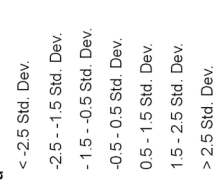
Leyenda



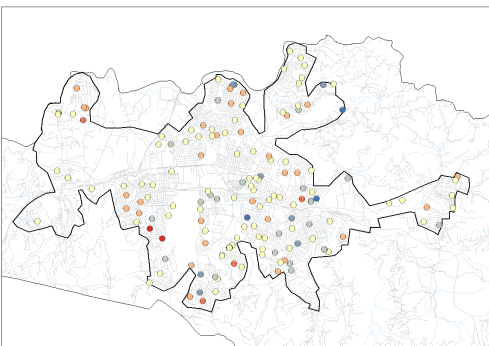
2007



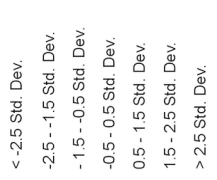
Leyenda



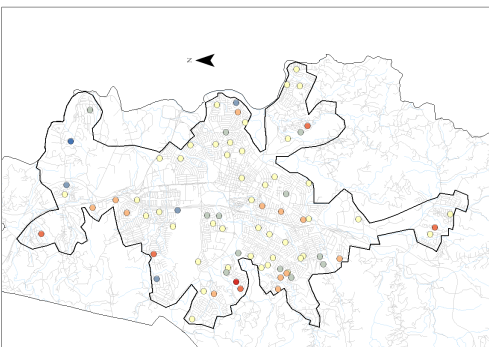
2008



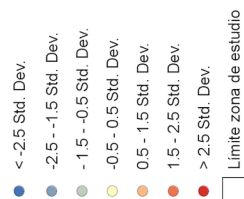
Leyenda



2009



Leyenda



6.2 Construcción del modelo con la técnica de Regresión Geográficamente Ponderada

En las diversas etapas de construcción del modelo utilizando el método de Mínimos Cuadrados Ordinarios se pudo observar la existencia de problemas de no estacionariedad espacial. Con el método de “Regresión Geográficamente Ponderada” (Geographic Weighted Regression) Brunson et al. (1996) buscan superar esta condición. Para ellos la no estacionariedad espacial es una condición en la que un modelo simple global no puede explicar las relaciones existentes entre algunos conjuntos de variables. Por tal razón la naturaleza del modelo a utilizar debe alterar el espacio donde se localizan las variables para poder reflejar la estructura interna de los datos. Con la Regresión Geográficamente Ponderada, Brunson et al. (1996) intentan captar esta variación, calibrando un modelo de regresión múltiple que permite la existencia de diferentes relaciones en diversos puntos en el espacio.

La calibración del modelo de Regresión Geográficamente Ponderada sigue un enfoque de mínimos cuadrados ponderados localmente (Fotheringham et al., 2002). A diferencia de los mínimos cuadrados ordinarios, al calibrar los coeficientes en la localización i , la Regresión Geográficamente Ponderada asigna pesos a través de un esquema de ponderación a los datos en las ubicaciones de acuerdo a su proximidad espacial a la posición i (Yu, 2007). Estos pesos actúan para asegurar que los lugares más cercanos ejercen una mayor influencia que aquellos lugares más alejados.

Se tomaron los mismos grupos de datos seleccionados para el modelo final utilizando el método de Mínimos Cuadrados Ordinarios. El método de Regresión Geográficamente Ponderada fue desarrollado utilizando el programa ArcGIS.

El primer paso en el análisis de los resultados fue evaluar los resultados de la autocorrelación espacial. Al evaluar la autocorrelación espacial dentro del desempeño del modelo, el valor obtenido de la prueba del Índice de Moran (Tabla 6.8) muestra que los valores del índice giran en torno al cero. Por este resultado se entiende que los datos presentan un patrón espacial totalmente aleatorio. En cuanto al valor Z , solo los años 2006 y 2007 muestran valores negativos. Esto indica que la observación está por debajo de la media. Los años restantes muestran valores positivos entendiéndose que las observaciones se encuentran sobre el valor de la media. Sin embargo, todos los valores observados se encuentran en el rango de

-1,65 a 1,65 lo que establece que estos son productos de una distribución aleatoria y no existe auto-correlación espacial.

TABLA 6.8: Índice Moran, Regresión Geográficamente Ponderada

	2005	2006	2007	2008	2009
Índice Moran	0,040299	-0,55404	-0,032723	0,037749	-0,002412
Valor Z	0,639262	-1,041088	-0,268875	0,497596	0,153693
P - valor	0,522653	0,297835	0,788026	0,618769	0,877852

Estos resultados fueron corroborados observando la distribución espacial de los residuales de la regresión. Se concluye que el modelo de regresión está bien especificado ya que se muestra una distribución aleatoria de sus residuales (Gráfica 6.5).

Por su parte, el Coeficiente de determinación múltiple R^2 promediado para cada año fluctuó para el período de estudio entre 0,88 y 0,93 (Tabla 6.8). En otras palabras, en este período de tiempo las variables independientes seleccionadas pueden explicar entre un 88% y un 93% el comportamiento de la variable dependiente. Es importante aclarar que la Regresión Geográficamente Ponderada, al ser un método local, genera Coeficientes de Determinación locales. Para entender la distribución de este coeficiente y el posible peso de variables en estos se crearon una serie de mapas (Gráficas 6:6 a 6:14).

Tabla 6.9: Regresión Geográficamente Ponderada - Coeficiente de Determinación Múltiple

Prueba	2005	2006	2007	2008	2009
Coeficiente determinación múltiple R^2	0,93	0,89	0,92	0,88	0,90

En cuanto al poder explicativo del modelo para el año 2005, los coeficientes de determinación R^2 locales muestran valores que fluctúan entre 85.07% y 93.94%. Se puede observar que los coeficientes mayores se concentran en la zona sur de la ciudad. En esta zona se localizan algunos de los asentamientos más recientes, en particular la urbanización Caguas Real. Esta urbanización se caracteriza por ser una de clase alta. Cuenta con control de acceso, piscina y centro de actividades. Además, es la única urbanización en Caguas que cuenta con

campo de golf. Para el año 2005 los valores residenciales en esta urbanización giraban entre los 650.000 y 900.000 dólares.

En el año 2006 la concentración del coeficiente de determinación R^2 locales más altos no solo incluyen las comunidades al sur de la ciudad, si no añaden comunidades en la franja oeste de la misma (Gráfica 6.4). Los coeficientes de determinación R^2 locales muestran valores que fluctúan entre 83,2% y 88,7%. Analizando las urbanizaciones que se encuentran en este renglón se observa que la mayoría corresponde a las que se han construido en los pasados 15 años. Representan los asentamientos más recientes y por tal razón responden a las tendencias recientes en el proceso de urbanización. En su inmensa mayoría cuentan con control de acceso, con 2 o 3 baños por residencia, amenidades como campo de béisbol o cancha de voleibol, y se encuentran muy cerca de centros comerciales. Responden a un mercado de clase media alta. Los valores residenciales en esta zona giraban para el 2006 entre los 250.000 y 400.000 dólares. Los coeficientes más bajos responden a urbanizaciones construidas en la década del 1950 al 1970. Acomodan a población de clase media baja y no cuentan con tantas amenidades como las que reflejan los coeficientes más altos.

El patrón observado en los años 2005 y 2006 se modifica en el año 2007 (Gráfica 6.4). En este año los valores del Coeficiente de determinación múltiple R^2 variaron entre 0,881 y 0,927. En otras palabras, explican entre 88,1% y 92,7% de la variable dependiente. Las expresiones más significativas se presentaron en pequeñas franjas del sur de la ciudad. Esto a diferencia de los años anteriores en donde se presentaban en gran parte de la zona sur y oeste de la ciudad. La zona norte de la ciudad presentó los coeficientes más bajos, entre 0,881 y 0,889. Los estudios plantean este año como el cual presentó la ruptura de la burbuja inmobiliaria. Los patrones de compra de inmuebles se vieron alterados de manera significativa. Las preferencias en amenidades en los complejos residenciales pierden la importancia que se observaba en años anteriores y las personas no están dispuestas a pagar las cantidades que años anteriores. Por otro lado, el mercado inmobiliario se ve invadido de residencias reposéidas⁹ por los bancos. Los bancos con el objetivo no acumular la cantidad de residencias bajo esta condición, colocan las mismas en el mercado de ventas a precios muy por debajo del valor original de las mismas. Los compradores aprovechan estas ofertas. A modo de ejemplo, la urbanización Caguas Real mencionada anteriormente por el campo de golf con

⁹ En Puerto Rico se utiliza el término “Residencia Reposeída” para las residencias embargadas por bancos o agencias de crédito debido a la falta de pagos. El término utilizado en inglés es “Foreclosure”. En otras regiones de Latinoamérica se le conoce como Remate o Ejecución Hipotecaria.

que cuenta, en un origen sus residencias tenían un costo entre los 650.000 y 900.000 dólares. Una vez surge la crisis inmobiliaria se llegaron a vender residencias por los bancos en 250.000 dólares.

En el año 2008 (Gráfica 6.4) se refleja un patrón parecido al de los años 2005 y 2006. Por otro lado, el año 2009 (Gráfica 6.4) muestra un patrón que se puede considera mezcla de los cuatro anteriores. En éste, la parte sur de la ciudad, el modelo logra explicar la variación mayor de la variable dependiente y el norte la variación es más baja. Para el año 2008 Coeficiente de determinación múltiple R^2 arrojó valores que fluctuaban entre 0,781 y 0,886 mientras que para el año 2009 estos valores fluctuaron entre 0,833 y 0,928.

La herramienta de análisis de Regresión Geográficamente Ponderada produce una serie de imágenes que permiten observar el comportamiento de los coeficientes a lo largo de la zona de estudio. Analizando éstas se puede determinar dónde la variable explicativa tuvo mayor relevancia en la predicción del valor del inmueble y en dónde tuvo menos. Con éstas se puede comprender mejor la variación en la zona de estudio de las variables explicativas. Al utilizar la técnica de Regresión Geográficamente Ponderada para modelar alguna variable no solo se busca predecir valores de la condición bajo estudio y los factores que influyen en esta predicción, también se busca comprender cuan espacialmente estacionarias son las relaciones entre la variable dependiente y las variables explicativas a lo largo de la zona de estudio. Examinando el resultado de los coeficientes y su distribución espacial permite observar donde se ubica ésta y cuánta variación está presente.

Al analizar la variable de la superficie del solar medido en metros cuadrados (Gráfica 6.6) se puede observar que para el año 2005 los coeficientes con valores más altos se concentran en la parte central de la ciudad. Esta zona corresponde a la más antigua de la ciudad. Por otro lado, los coeficientes con valores más bajos se pueden observar en la parte sur de la ciudad. Los coeficientes con valores más altos indican un valor aproximado de \$67 por metro cuadrado. En las zonas con valores más bajos el coeficiente muestra un valor que gira en torno a los \$42 por metro cuadrado. Ya en el año 2006 los coeficientes con valores más altos se concentran en el centro de la ciudad con dirección hacia la zona este de la misma. Para esta zona el coeficiente indica un valor de \$23 por metro cuadrado. Esto es una diferencia alrededor de \$44 cuando se compara con el año 2005. En cuanto a la zona con el valor de coeficiente más bajo la misma se muestra al norte de la ciudad con precios aproximados de \$5 por metro cuadrado. Una diferencia de unos \$37 cuando se compara con el año anterior.

En el año 2007 se reflejan los valores mayores por metro cuadrado. La zona donde se muestran los valores mayores parten del centro de la ciudad en dirección o este de la misma. En esta los valores por metro cuadrado giran en torno a los \$187. Las zonas más bajas se observan al sur este de la ciudad con valores que giran en torno a los \$131. El año 2008 muestra un patrón similar al del 2007 sin embargo los valores por metro cuadrado son muy diferentes. En este año los valores más altos giran en torno a los \$121 llegando a ser el más bajo unos \$7. Ya para el año 2009 la zona con los coeficientes de mayor valor se localiza en toda la parte suroeste de la ciudad. En ésta el costo del metro cuadrado se establece en torno a los \$164 llegando hasta la zona más baja que se concentra en la parte noreste de la ciudad en donde el costo del metro cuadrado se ubicó en torno a los \$87. Morfológicamente, la ciudad presenta los solares de menor tamaño en el centro de la ciudad. Coinciden los modelos de los 5 años en localizar en esta zona la mayor influencia de la variable superficie del solar medido en metros cuadrados. Se puede entender que esta influencia mayor en esta zona se debe a la localización de los inmuebles en la parte central de la ciudad.

En cuanto a la variable edad (Gráfica 6.7) se asume que por cada año de construcción que tenga el inmueble su valor se reduce. Se considera una variable con influencia negativa. Para el año 2005 se observa que esa reducción fue más marcada en el sur de la ciudad. En esta zona un inmueble podría reducir su valor hasta \$1.463 por cada año de construido. La zona donde más valor perdían los inmuebles por esta variable se concentra en el norte de la ciudad. En ésta un inmueble podría perder hasta \$2.000 por cada año de construcción. Observando el año 2006, es en la zona suroeste de la ciudad donde menos valor pierden los inmuebles por condición de su edad. La pérdida puede llegar a ser de unos \$1.268 por cada año de existencia del inmueble. En el lado opuesto el inmueble puede llegar a perder hasta \$1.622 de su valor por cada año del inmueble. Esta condición se observa mayormente en la parte este de la ciudad. Ya en el año 2007 la mayor pérdida se puede observar al oeste de la ciudad. En esta un inmueble puede reducirse \$2.354 por cada año de construcción. Donde menos pierde el inmueble valor por esta condición es al este de la ciudad donde la pérdida puede llegar a ser \$1.622 por año. Los años 2008 y 2009 muestran patrones similares. La mayor pérdida en valores se observa en la zona noroeste de la ciudad. Para el año 2008 esta pérdida significaba \$1520 por cada año y para el año 2009 la cifra giraba en torno a los \$2.070. En cuanto a la menor pérdida de valor en este renglón en el año 2008 la cantidad era de \$921 por cada año y en el año 2009 una cantidad de \$1.606. La distribución de los coeficientes de la variable edad en la zona de estudio responde a la existencia de nuevos complejos residenciales en la ciudad.

En cada año analizado se observa que en la zona donde se refleja la concentración de coeficientes con valores más altos coincide con la localización de nuevos inmuebles.

Al observar la variable distancia a centro comercial (Gráfica 6.8) se puede observar en los años 2005, 2006 y 2007 muestran un patrón similar. Al igual que la variable edad del inmueble, esta variable se considera una influencia negativa. En otras palabras, cuanto mayor distancia separe el inmueble del centro comercial, menor es su valor. Al computar el promedio de los coeficientes para cada año se observó que en efecto la influencia era negativa. Sin embargo, al observar las distribuciones locales de los coeficientes y su influencia se puede constatar que en algunas zonas su influencia fue positiva. Esta situación se reflejó en los años 2005, 2006 y 2007. Para estos tres años la menor pérdida de valor del inmueble por estas condiciones se observó en el norte de la ciudad. Se puede asumir que por la gran concentración de centros comerciales en esta zona existe una menor distancia de traslado entre las urbanizaciones y los centros comerciales reflejándose esto en una pérdida menor de valor por esta condición. Ya en el año 2008 se observa un cambio en el patrón en el cual la menor pérdida se refleja en la zona oeste de la ciudad. Esta zona también coincide con la existencia de centros comerciales. Además, muestra la construcción durante ese período de complejos de megatiendas como es el caso de Walmart. La pérdida de valor es alrededor de unos ocho dólares en el año 2009. Este se refleja en pequeña zona del este del término municipal.

La superficie de ocupación del inmueble medida en metros cuadrados es una variable de gran importancia en el modelo de valoración (Gráfica 6.9). Durante los cinco años bajo análisis el año 2006 mostró el mayor valor por metro cuadrado que podía adquirir un inmueble por su superficie de ocupación. Esta cifra fue de \$1.630 por metro cuadrado de superficie del inmueble. La cantidad menor se observó en el año 2007 con una cifra de \$439 por metro cuadrado de superficie de ocupación del inmueble. El año que mayor valor mostró esta variable fue el año 2006 en el cual el costo por superficie de ocupación giro entre los \$1.630 y los \$1.249 por metro cuadrado. El año con el desempeño más bajo en este renglón fue el 2009 en el que los valores giraban entre los \$570 y los \$441 por metro cuadrado. En los cinco años se pudo observar que la mayor influencia en este renglón se concentraba en la zona norte de la ciudad. Es importante destacar que en esta zona se construyeron urbanizaciones de carácter exclusivo con viviendas que cuentan con superficies de ocupación superiores siendo en ocasiones el doble de lo que ocupa una residencia de clase media o baja en el centro de la ciudad. Las zonas con menor influencia en este renglón se localizaron en el centro de la

ciudad discurriendo hacia el sur de la misma. El año 2009 refleja una excepción a este patrón al observar que la zona con mayor influencia se extendía al sur de la ciudad.

Otra de las variables bajo análisis fue la existencia de cancha de tenis (Gráfica 6.10). Los años 2005 y 2006 reflejan patrones que van de acuerdo a la estructura de la ciudad. En estos años las urbanizaciones localizadas al norte de la ciudad eran las que incluyan dentro de sus complejos canchas de tenis. Se puede deducir que por esta razón se observa un patrón de influencia mayor de esta variable en esa zona. En el año 2007 se inauguran y entran en mercado mobiliario urbanizaciones con grandes complejos deportivos en particular con varias canchas de tenis. Éstas se localizan en la zona central-oeste de la ciudad. Por esto se observa cómo en este año la mayor influencia por esta variable se traslada a esta zona. Este año muestra la mayor influencia en el valor del inmueble por esta variable con la cifra de \$61.759 por la existencia de cancha de tenis dentro del complejo residencial. La influencia menor por esta variable se presentó en el año 2008 en la cual la existencia de cancha de tenis implicaba un aumento de \$6.400 al inmueble.

La variable gimnasio, (Gráfica 6.11) tiene un comportamiento similar en los años 2005, 2006, 2008 y 2009. En todos ellos la mayor influencia de la variable se presenta al sur de la ciudad. Esto coincide con las urbanizaciones completadas para la fecha del estudio. Estas en su mayoría incluyen gimnasio como parte de las amenidades de la urbanización. En un patrón diferente se muestra el año 2007. En este no se evidencia que la ausencia de gimnasio significara una pérdida de valor para el inmueble. Los valores para este año fluctuaban entre los \$47.428 y los \$31.929. Ya se ha observado en el comportamiento de diversas variables para este año (2007) patrones atípicos a los reflejados en los años previos o posteriores. También se indicó anteriormente que este es el año en que se establece la ruptura de la burbuja inmobiliaria. Se puede adjudicar ese comportamiento atípico al impacto de la ruptura de la burbuja en el patrón de adquisición de inmuebles.

En cuanto a la variable cancha de voleibol (Gráfica 6.12) se puede observar que los años 2005, 2006, y 2008 muestran patrones muy similares en cuanto al comportamiento de los coeficientes a lo largo de la zona de estudio. En los tres años los coeficientes más elevados se reflejan al oeste de la ciudad. Esta influencia fluctúa entre los \$75.000 y los \$27.000. El año 2007 presenta un patrón en el cual los mayores coeficientes se ubican en la parte norte de la ciudad. El año 2009 presenta un patrón parecido al del año 2007, solo que se observa mayor influencia en la zona norte. En los cinco años se puede observar que la menor influencia, o

los coeficientes más bajos, se alojan en la parte este de la ciudad. Esto coincide con las comunidades más antiguas de la ciudad las cuales en su mayoría no cuentan con canchas de voleibol.

La variable campo de béisbol (Gráfica 6.13) esta presenta un patrón extremadamente parecido en los primeros cuatro años del estudio variando el mismo en el año 2009. En estos primeros años los coeficientes más bajos se localizan de manera concentrada al este de la ciudad cambiando totalmente en el año 2009 al lado oeste de la ciudad. Los coeficientes de mayor desempeño se observan en la parte central oeste de la ciudad. Muchas de las urbanizaciones de esa zona se construyeron en las décadas del 1960 y 1970. Era muy común que las urbanizaciones incluyeran parque de béisbol dentro de sus amenidades. Los valores de los coeficientes para esta variable pueden llegar a la cifra de \$85.500 para el año 2007, hasta la cantidad de \$14.168 en el año 2005.

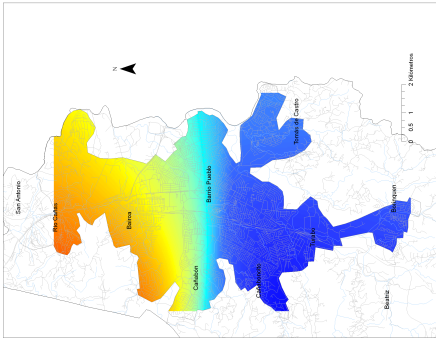
Finalmente, la variable campo de golf (Gráfica 6.14) presenta la influencia que podríamos asumir dada la localización del único campo de golf en la ciudad. El mismo se encuentra localizado en la esquina suroeste de la ciudad y por tal razón se pueda observar que en la mayor influencia de esta variable se localiza en esa zona. Sin embargo, de todos los años de estudio es sólo el año 2009 el que no presenta un patrón similar al de los años anteriores. En éste, la mayor influencia de la variable se concentra en la zona este de la ciudad. Dejando de esa manera la zona suroeste como una de baja influencia en los coeficientes. Un inmueble localizado dentro de una comunidad con campo de golf puede llegar a apreciar en su valor la cantidad de \$56.190, esto para el año 2009, hasta la cantidad de \$14.663 esto para el año 2007.

Gráfica 6.5: Residuales de Regresión, Regresión Geográficamente Ponderada para los años 2005 al 2009

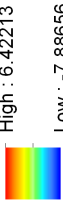


Gráfica 6.8: Distribución Coeficiente de Variable CC_Dist (Distancia centro comercial) - RGP Años 2005 al 2009

2005

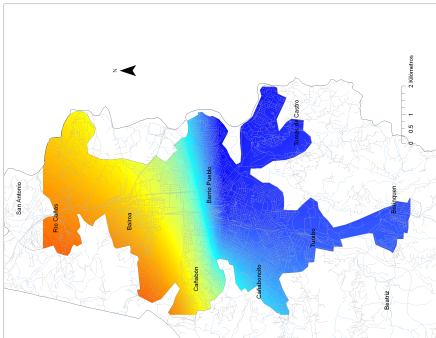


Leyenda



High : 6.42213
Low : -7.88656

2006

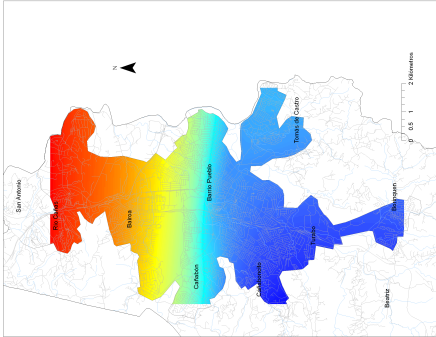


Leyenda



High : 7.8084
Low : -1.56082

2007

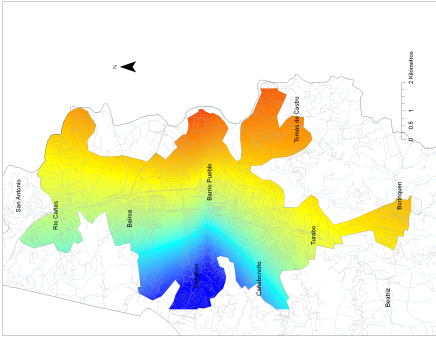


Leyenda



High : 3.18023
Low : -8.69168

2008

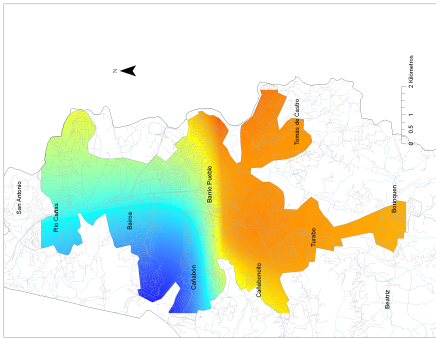


Leyenda



High : -0.206398
Low : -12.2176

2009



Leyenda

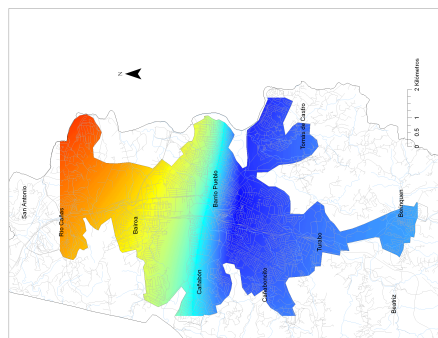


High : -8.13742
Low : -17.2971

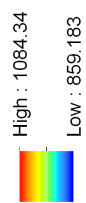
Gráfica 6.9: Distribución Coeficiente de Variable HouseAreaM (Superficie de ocupación en M²) - RGP

Años 2005 al 2009

2005

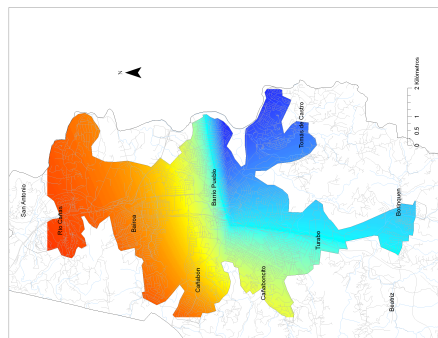


Leyenda



High : 1084.34
Low : 859.183

2006

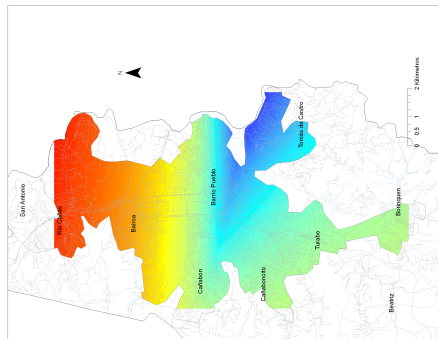


Leyenda



High : 1630.98
Low : 1249.08

2007

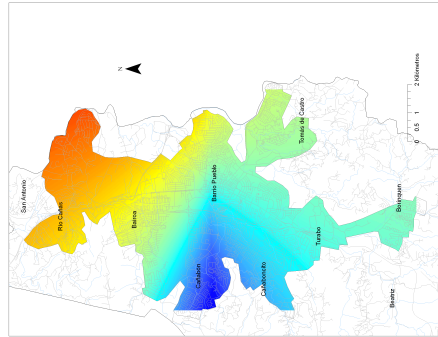


Leyenda



High : 689.967
Low : 439.802

2008

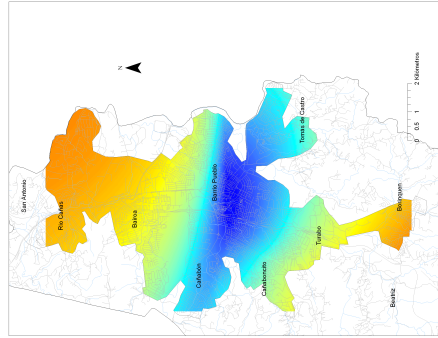


Leyenda



High : 1354.21
Low : 930.401

2009



Leyenda

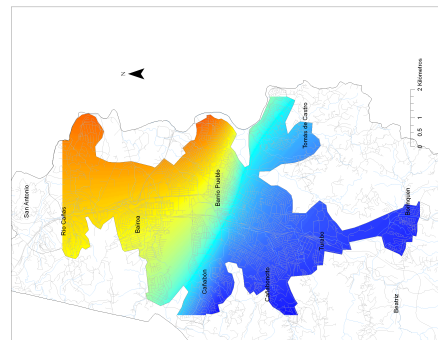


High : 570.154
Low : 441.024

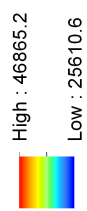
Gráfica 6.10: Distribución Coeficiente de Variable Tenis (Existencia cancha de tenis) - RGP

Años 2005 al 2009

2005

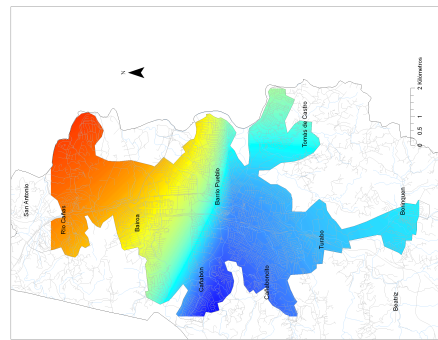


Leyenda



High : 46865.2
Low : 25610.6

2006

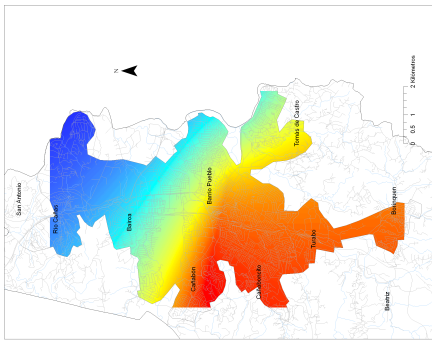


Leyenda



High : 34612.7
Low : 17933

2007

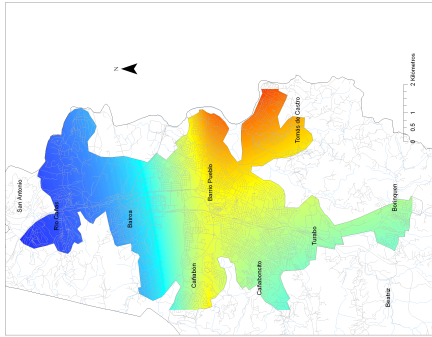


Leyenda



High : 61759.6
Low : 31852.3

2008

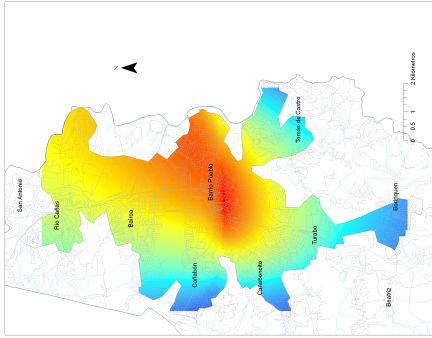


Leyenda



High : 20082.3
Low : 6400.05

2009



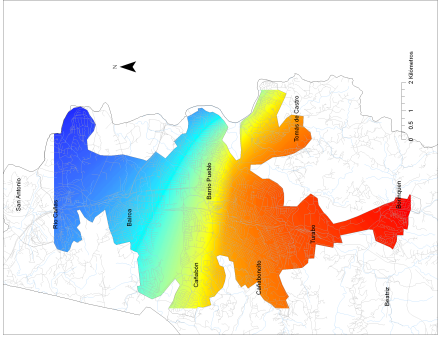
Leyenda



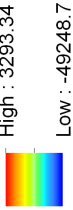
High : 35963
Low : 30624

Gráfica 6.11: Distribución Coeficiente de Variable GYM (Existencia de facilidades de gimnasio para ejercicios) - RGP Años 2005 al 2009

2005

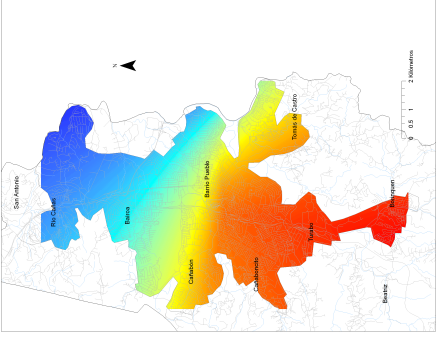


Leyenda

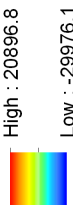


High : 3293.34
Low : -49248.7

2006

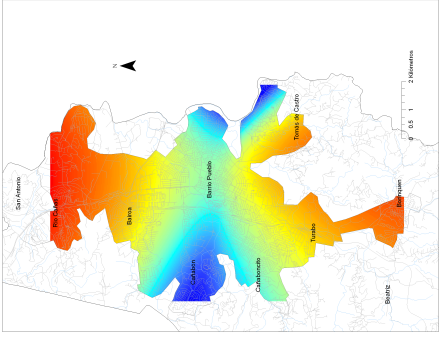


Leyenda

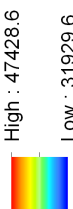


High : 20896.8
Low : -29976.1

2007

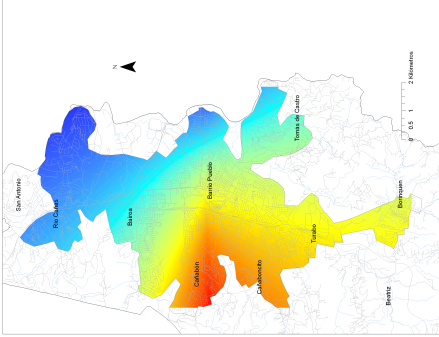


Leyenda

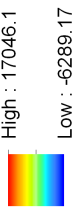


High : 47428.6
Low : 31929.6

2008

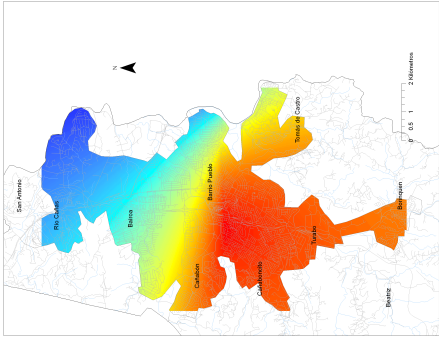


Leyenda

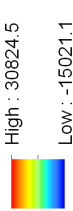


High : 17046.1
Low : -6289.17

2009



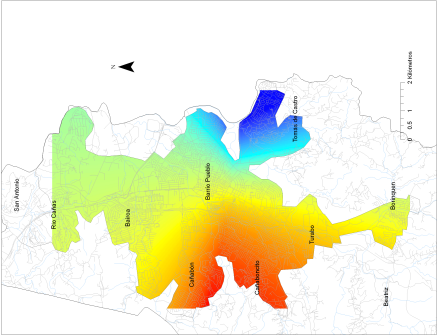
Leyenda



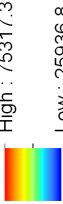
High : 30824.5
Low : -15021.1

Gráfica 6.12: Distribución Coeficiente de Variable VB (Existencia de cancha de voleibol) - RGP Años 2005 al 2009

2005

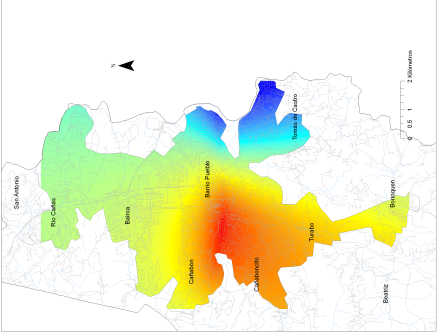


Leyenda



High : 75317.3
Low : 25936.8

2006

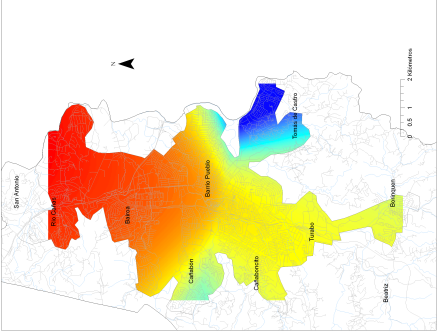


Leyenda



High : 27255.7
Low : -4312.94

2007

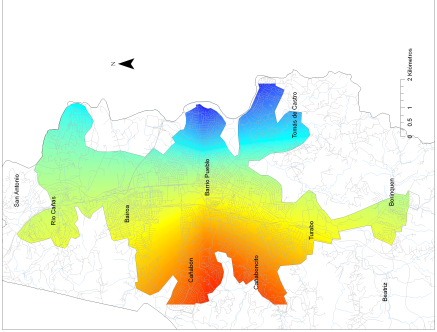


Leyenda



High : 13331
Low : -7835.42

2008

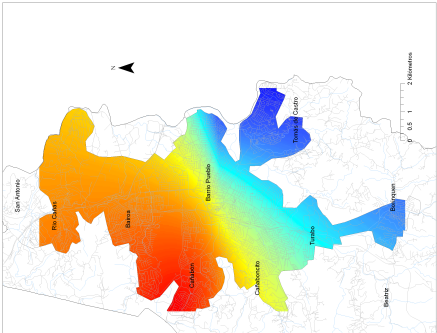


Leyenda



High : 70259.9
Low : 24822

2009



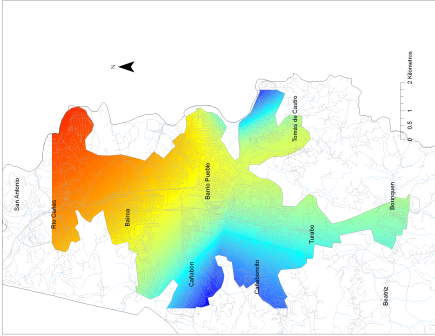
Leyenda



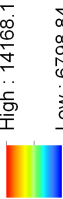
High : 56323
Low : 20671.4

Gráfica 6.13: Distribución Coeficiente de Variable BB (Existencia de parque de béisbol) - RGP Años 2005 al 2009

2005

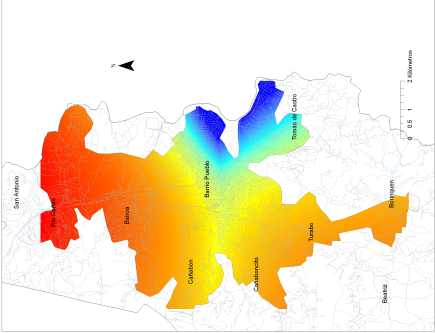


Leyenda



High : 14168.1
Low : 6798.84

2006

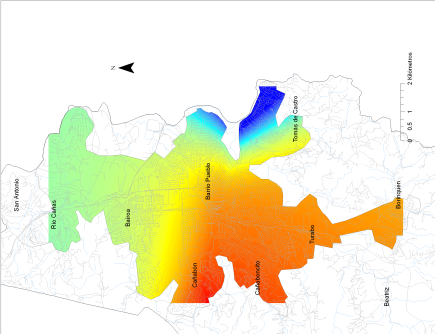


Leyenda



High : 30706.9
Low : 13082.4

2007

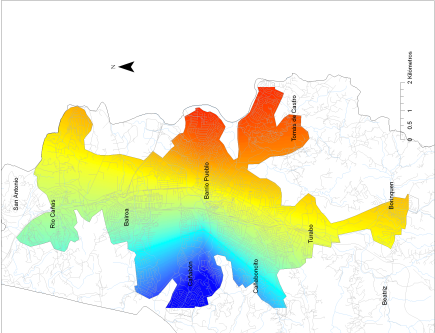


Leyenda



High : 8515.51
Low : -2463.05

2008

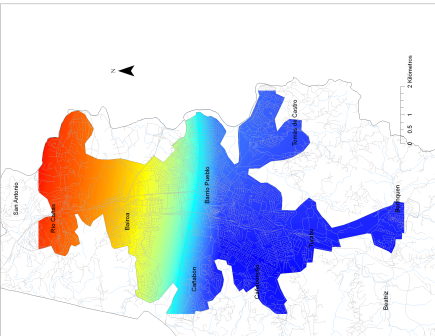


Leyenda



High : 17298.5
Low : -915.827

2009



Leyenda



High : 17885.5
Low : -10597.3

6.3 Comparación de Modelos

Con el objetivo de realizar la selección del modelo que mejor representara la predicción de los valores inmuebles de los años bajo estudio se realizaron una serie de pruebas estadísticas. La primera de ellas fue la del error cuadrático medio. Está es una medida de dispersión del error del pronóstico, o, dicho de otra forma, mide el tamaño del error en unidades. Para ambos modelos se aplicó esta técnica y se obtuvo una serie de resultados que se compararon año a año (Tabla 6.10). Esta prueba refleja que la regresión geográficamente ponderada mostró un mejor desempeño cuando se compara con la de los mínimos cuadrados ordinarios. En promedio la diferencia entre los años 2005 y 2008 fue de unos \$2.500. En el caso del año 2009 la diferencia fue mucho más marcada. Esta fue de \$21.719.

TABLA 6.10: Comparación Técnicas – Error Cuadrático Medio (dólares)

Año	Mínimos Cuadrados Ordinarios	Regresión Geográficamente Ponderada
2005	44.522	42.345
2006	54.865	52.040
2007	46.957	44.846
2008	43.496	41.232
2009	69.331	47.612

Por otro lado, se calculó el Error Absoluto Medio para cada año y cada modelo y se procedió a comparar los resultados de los mismos (Tabla 6.11). El Error Absoluto Medio es una medida robusta de la variabilidad de una muestra de datos cuantitativos. Nuevamente se observa como la regresión geográficamente ponderada muestra un mejor desempeño cuando se compara con la de los mínimos cuadrados ordinarios. Exceptuando el año 2007 se puede establecer que el promedio de diferencia entre los dos modelos es de aproximadamente \$2500. El año 2007 es en el que menor diferencia estando el valor de la regresión geográficamente ponderada tan sólo \$301 por debajo de la del valor arrojado con la técnica de los mínimos cuadrados ordinarios.

TABLA 6.11: Comparación Técnicas - Error Absoluto Medio (MAE) (Dólares)

Año	Mínimos Cuadrados Ordinario (MCO)	Regresión Geográficamente Ponderada (RGP)
2005	32.935	30.219
2006	37.324	35.470
2007	30.328	30.027
2008	30.780	28.056
2009	32.523	29.752

Como tercer paso se procedió a comparar el coeficiente de determinación múltiple R^2 y la prueba de Akaike (AICc) (Tabla 6.12). En los cinco años de análisis se pudo observar que los valores de los coeficientes de determinación múltiple obtenidos de la prueba de regresión geográficamente ponderada demostraban mejor desempeño cuando se comparaban con el de la técnica de los mínimos cuadrados ordinarios. Esto implica que bajo la técnica de regresión geográficamente ponderada las variables independientes explican mejor el valor de la variable dependiente. Se entiende que el modelo con el valor Akaike más bajo es el que mejor se ajusta a los datos observados. Esto tomando en cuenta la complejidad del modelo. Se puede observar que los cinco años bajo estudio esta prueba arroja valores más bajos cuando se ejecutó para la regresión geográficamente ponderada.

TABLA 6.12: Comparación Técnicas – Resultados R2 y AICc

	2005		2006		2007		2008		2009	
	MCO	RGP	MCO	RGP	MCO	RGP	MCO	RGP	MCO	RGP
R2	0,91	0,93	0,88	0,89	0,90	0,92	0,85	0,88	0,86	0,90
AICc	2517	2502	3131	3124	2122	2117	3799	3772	1875	1863

6.4 Selección de Modelo

Observada esta serie de pruebas y tomando en consideración sus resultados se concluye que la técnica de la Regresión Geográficamente Ponderada presenta un mejor desempeño en la predicción de los valores de los inmuebles residenciales cuando se compara la misma con la técnica de los mínimos cuadrados ordinarios.

Capítulo 7

Validación del Modelo

7. Validación del Modelo

Al comenzar esta investigación se planteó que los criterios que se utilizan para la valoración y cobro de contribuciones sobre los inmuebles en Puerto Rico no permiten que el sistema sea justo ni equitativo ya que estos datan de un procedimiento establecido en 1951 y ajustado en 1957, esto es, hace más de cincuenta años, sin revisión de los mismos. Además, dado que no se realiza un proceso de actualización de los datos y los dueños de inmuebles no informan de cambios estructurales en términos de mejoras como ampliaciones, construcción de terrazas o de piscinas, estos cambios estructurales del inmueble no están contemplados en el valor de tasación asignado. Se tributa basado en el valor asignado originalmente al inmueble una vez se inscribió y se tasó por primera vez con las tasas establecidas en 1957.

Con el objetivo de establecer un método que permitiera modernizar el sistema de valoración inmueble se decidió adoptar una técnica basada en la predicción de datos de ventas residenciales efectuadas en tiempo reciente. Para esto se compararon las técnicas de valoración hedónica utilizando los Mínimos Cuadrados Ordinarios y la técnica de Regresión Geográficamente Ponderada. Utilizando Sistemas de Información Geográfica se procesaron datos de ventas residenciales en un período de cinco años (2005 al 2009) y se aplicaron estas dos técnicas. De los resultados obtenidos se concluyó que la Regresión Geográficamente Ponderada permite realizar una predicción más certera de los datos.

Con el análisis de los modelos también se busca estudiar el efecto de eventos particulares en los procesos de venta de inmueble. En este caso se tomaron los datos de ventas residenciales entre los años 2005 y 2009 porque se buscaba analizar el efecto que podría tener el estallido de la burbuja inmobiliaria en que se encontraba encerrado el mercado inmobiliario de Puerto Rico. Los expertos plantean que el momento de la ruptura sucede en el año 2007 (Coy, 2008; Krugman, 2009; Shiller, 2011).

Con el objetivo de estudiar la vigencia a largo plazo de los valores de predicción obtenidos por los modelos se procedió a obtener datos de ventas residenciales para los cinco años posteriores a los seleccionados para el análisis de los modelos; esto es, datos de ventas residenciales para la zona de estudio entre los años 2010 y 2014. Se tomaron los datos de predicción de los años 2005 al 2009 y se asumió que estos serían las predicciones de valor inmueble de referencia para los años subsiguientes. Se identificó el período de tiempo que estas predicciones mantenían su vigencia.

El estudio de la vigencia de los valores de predicción se fundamentó en tres pruebas. La primera es el Coeficiente de Dispersión. El Coeficiente de Dispersión (COD) es una medida de la uniformidad de la valoración de todas las propiedades de una ciudad. Mide la desviación media entre los precios de venta de las propiedades vendidas de la media de valoración de inmuebles de la ciudad. En otras palabras, si hay dos inmuebles en una ciudad y cada uno se evalúa por la misma cantidad, el COD mide la probabilidad de que vendan por la misma cantidad. Se entiende que un COD de 0,00% sería el valor perfecto, ya que indica absoluta equidad en la medida en que cada inmueble es valorado exactamente en el mismo porcentaje del valor justo de mercado. Sin embargo, un COD de 0.00% se considera inalcanzable.

La Asociación Internacional de Oficiales Evaluadores (International Association of Assessing Officers o IAAO por sus siglas en inglés) publica normas de uniformidad utilizando el coeficiente de dispersión. El COD se calcula tomando la diferencia entre la razón de predicción y valor de venta de cada propiedad y la razón mediana de predicción y valor de venta (ignorando los signos positivo y negativo), sumando estas diferencias y dividiendo por el número de propiedades. Esto determina la desviación media de la mediana. Esta cantidad se divide por la mediana para expresar el resultado como un porcentaje de la mediana. Por ejemplo, un COD de 15 por ciento significa que las propiedades tienen relaciones que, en promedio, se desvían un 15 por ciento de la razón mediana.

El COD y la desviación porcentual mediana se calculan de manera similar. Sin embargo, la desviación porcentual mediana utiliza la desviación mediana mientras que el COD utiliza la desviación media. Al calcular la desviación porcentual media sólo importa si la razón entre la predicción y el valor de venta de una propiedad está sobre o bajo la mediana. No importa qué tan lejos esté sobre o bajo de la mediana. Pero al calcular la desviación media, lo que importa es la cantidad de la razón entre la predicción y el valor de venta de la propiedad está sobre o bajo la mediana. Las relaciones muy alejadas sobre o bajo la mediana tienen más influencia que las propiedades con relaciones cercanas a la mediana. Esto significa que el COD tiende a ser mayor que la desviación porcentual media.

Como regla general la IAAO establece que un COD de un 15% o menos refleja un nivel relativamente alto de equidad entre los valores de venta y la predicción de valor de los inmuebles. Esto es solo una sugerencia de la organización ya que cada territorio tiene la

libertad de adoptar el valor de COD aceptable. A modo de ejemplo, en el estado de Vermont en los Estados Unidos, el COD debe ser de 20% o menos.

La utilización del COD como medida de uniformidad es una aceptada por los especialistas en valoración de inmuebles para fines contributivos. Sin embargo, el proceso para computar el mismo no considera tres factores que afectan su resultado: el valor límite adoptado, los niveles de confianza y el tamaño de la muestra (Gloudemans, 1999 y 2001). En particular el tamaño de la muestra seleccionada puede provocar que los valores del COD presenten cifras que superen los estándares adoptados. En cuanto al uso de intervalos de confianza, el IAAO no ha producido fórmulas que permitan calcular éstos para el COD (Gloudemans, 2001; Benmamqun, 2006). Para superar esta deficiencia Robert Gloudemans (2001) ha propuesto una fórmula que permite computar los límites máximos aceptables del COD fundamentado en los tres factores. Con ésta se busca, adoptando un valor COD como referencia, hasta qué punto pueden ser aceptados unos valores de COD superiores al de referencia tomando en consideración la diferencia en tamaño de la muestra y los niveles de confianza de cada muestra.

La fórmula se esboza de la siguiente manera:

$$MAXCOD = CODSTD \times \sqrt{\frac{x^2}{n-1}}$$

donde:

- MAXCOD es el valor máximo del COD que puede ser aceptado basado en el tamaño de la muestra
- CODSTD es valor límite adoptado para el COD
- (n-1) son los grados de libertad.

Para computar esto se tomó como referencia el trabajo realizado por Benmamqun (2006). Se tomó la rutina programada dentro del programa SPSS y se adaptó para el caso bajo estudio. La rutina requiere la utilización de una base de datos que incluye para cada inmueble el valor de predicción, el valor de venta y la razón entre estos dos valores. En la rutina se estableció como valor del COD deseado el 15% (el sugerido por la IAAO) a un nivel de confianza del 95%. Se procedió a calcular el COD máximo aceptable para cada año bajo análisis utilizando los valores de predicción obtenidos de los modelos de regresión geográficamente ponderada (Ver Tabla 7.1). De utilizar estrictamente el límite de 15% calculado según los procesos

establecidos por el IAAO, serían solo los años 2005, 2006, 2008 y 2009 los que cumplen con éste al obtener un COD de 12,24, 14,98, 13,35 y 14,57, respectivamente. Con este criterio el año 2007 no cumpliría con el límite adoptado. Sin embargo al observar el valor máximo del COD ajustado por el tamaño de la muestra y los límites de confianza, el COD máximo a ser aceptado para una muestra compuesta de 213 inmuebles es de 16,18. Bajo un modelo de predicción en masa que modela la predicción en intervalos anuales, se puede concluir que los modelos de los cinco años cumplen con la regla general sugerida por la IAAO cuando se computa el nivel máximo aceptable del COD tomando en consideración la diferencia en tamaño de la muestra y los niveles de confianza de éstas.

Se ha establecido que eventos extremos en el mercado inmobiliario pueden generar influencia en el proceso de valoración (Bowman et al., 1986; Stevenson, 2008). En esta investigación el año a observar es el 2007 debido al estallido de la burbuja inmobiliaria que atravesaba el mercado inmobiliario de Puerto Rico en la época. El comportamiento irregular de ventas de inmuebles observado para este año afecta de alguna manera el cálculo del COD.

TABLA 7.1: Valores Coeficiente de Dispersión (COD) obtenidos y valores COD máximos aceptables.

	V2005	V2006	V2007	V2008	V2009	V2010	V2011	V2012	V2013	V2014
<i>N</i>	467	816	213	435	212	153	215	281	266	222
<i>MAXCOD</i>	15,80	15,60	16,18	15,83	16,19	16,40	16,18	16,03	16,06	16,16
Pred05	12,24	13,56	14,26	14,37	15,59	12,61	14,53	16,74	18,02	17,52
Pred06		14,98	17,56	15,39	15,80	15,46	19,39	18,94	19,78	20,48
Pred07			16,04	15,18	15,79	14,27	15,39	17,64	18,75	19,47
Pred08				13,35	15,48	13,43	16,57	17,06	18,51	18,79
Pred09					14,57	14,54	15,39	16,42	18,33	17,54

Observando las predicciones, y bajo la idea de utilizar las mismas para el establecimiento del valor inmueble en períodos mayores al del año en que se efectuaron, se puede concluir que el mejor desempeño en este renglón lo presentó el año 2005. Con el valor de predicción de este año se pudieron obtener valores que cumplieron con el COD máximo aceptable entre los años 2005 al 2011. A éste le siguió las predicciones del año 2007 que mostraron poder ser utilizadas por cinco años. Las predicciones de los años 2008 y 2009 mostraron que pudieron ser utilizadas por tres años corridos. El efecto del estallido de la burbuja inmobiliaria se puede relacionar a la interrupción que se ve en el año 2007 cuando se utilizan los datos de

predicción del año 2006. El cumplimiento con el COD máximo aceptable se ve interrumpido en este año y se puede suponer que esto es debido al comportamiento errático de las ventas de inmuebles en ese año. La incertidumbre del mercado motivo a muchos dueños de inmuebles a vender sus residencias muy por debajo del valor de venta de mercado, mientras otros ignoraban la situación del mercado y conseguían compradores de inmuebles que adquirían los inmuebles a precios inflados (Marxuach, 2010). Con la excepción del uso del valor de predicción del año 2006 con relación a los valores inmuebles del 2007, todos los valores de COD obtenidos cumplen con el COD máximo esperado de cada año hasta llegar al año 2010. A partir del año 2012 el valor de predicción pierde efectividad ya que los COD presentan valores que no cumplen con el COD máximo aceptable.

En resumen, basado en los resultados del COD obtenido para cada año y el COD máximo esperado, la utilización de las predicciones de inmuebles se puede utilizar para el proceso de establecer las tasas a cobrar por medio de impuestos al inmueble, por un término máximo de tres años. Esto es cónsono con los modelos estudiados en otros países de la región (Commeaux, 2015; Montaña, 2016; de Cesare, 2016). La situación del valor de predicción del año 2006 con relación a los valores inmuebles del 2007 se toma como una particular que se esperaba sucediera.

Para entender mejor la dimensión de las diferencias entre los valores de predicción para cada año y cómo varía su influencia en el tiempo se realizaron otras pruebas. Una de estas fue la del Error Porcentual Absoluto Medio (MAPE o Mean Absolute Percentage Error). El Error Porcentual Absoluto Medio es un indicador del desempeño del pronóstico de demanda que mide el tamaño del error (absoluto) en términos porcentuales. El hecho que se estime una magnitud del error porcentual lo hace un indicador frecuentemente utilizado por los encargados de elaborar pronósticos debido a su fácil interpretación.

El error más bajo ocurrió con los valores de predicción del año 2005 y las ventas del año 2005 (tabla 7.2). Este error fue de un 15,18%. Utilizando como base el valor de predicción del año 2005 se nota que este error aumenta de manera incremental hasta ver un error de un 58,0% para los valores de venta del año 2014. Una revisión de los valores de predicción y el error en los años en que se tomó la muestra refleja que el error porcentual absoluto medio varía entre un 15,18% hasta un 22,12%. Observando el desempeño temporal de los valores de predicción se observa que el que mejor desempeño presentó a largo plazo y fue el del año 2009 cuando el mayor error fue de un 47,25% para los valores de venta del año 2014. El que

mayor error presentó fue el valor de predicción del año 2007 cuando se utiliza como referencia para las ventas del año 2014. En éste se observó un error porcentual absoluto medio de 59, 17%.

TABLA 7.2: Error Porcentual Absoluto Medio (MAPE) – Regresión Ponderada Distancia

	V2005	V2006	V2007	V2008	V2009	V2010	V2011	V2012	V2013	V2014
Pred05	15,18%	18,29%	20,66%	21,14%	27,48%	26,57%	33,39%	38,45%	40,27%	58,50%
Pred06		19,20%	20,06%	21,65%	25,40%	24,58%	31,96%	37,51%	40,22%	58,39%
Pred07			20,23%	22,81%	29,15%	25,64%	33,31%	38,69%	41,30%	59,17%
Pred08				16,78%	21,46%	20,07%	26,84%	31,69%	33,97%	50,62%
Pred09					22,12%	19,20%	25,63%	29,17%	31,60%	47,25%

Además del Error Porcentual Absoluto Medio se procedió a calcular el Error Absoluto Medio (MAE). El Error Porcentual Absoluto Medio da una idea clara del desempeño de los datos analizados, pero su significado en términos monetarios da una idea más clara del impacto en dólares que presentan estas relaciones. La menor diferencia entre la relación de predicción cuando se compara con el valor de venta se observa en la relación del valor de predicción con datos del año 2008 y los valores de venta de inmuebles para el año 2010 (tabla 7.3). En términos económicos esta diferencia fue de tan solo \$21,693. Cuando se compara los valores de predicción generados versus los cinco años en que obtuvieron estos se observa que la mayor diferencia fue en los valores de predicción del año 2006 cuando se compara con los valores de venta de ese mismo año. La diferencia que muestra la predicción es de \$35,470. Las cifras para estos cinco años de predicción giran en torno a los \$31,000. Se asume que esta cifra está cerca de un 20% del valor promedio de los inmuebles para la zona de estudio. El año que mayor diferencia muestra es el de la predicción del año 2007 cuando se utiliza como base para los valores del año 2014. Esta diferencia es de \$69,278. Esta cifra es cerca del 50% del valor promedio de las viviendas en la zona de estudio.

TABLA 7.3: Error Absoluto Medio (MAE) - Regresión Ponderada Distancia

	V2005	V2006	V2007	V2008	V2009	V2010	V2011	V2012	V2013	V2014
Pred05	\$30.219	\$33.953	\$31.126	\$35.817	\$37.090	\$29.032	\$39.219	\$45.450	\$57.901	\$65.515
Pred06		\$35.470	\$30.634	\$38.357	\$35.756	\$27.555	\$39.091	\$46.923	\$58.573	\$67.259
Pred07			\$30.027	\$36.479	\$39.407	\$29.179	\$39.933	\$46.735	\$60.038	\$69.278
Pred08				\$28.056	\$29.596	\$21.693	\$31.631	\$37.970	\$48.874	\$57.243
Pred09					\$29.752	\$23.598	\$30.059	\$33.475	\$43.645	\$52.028

Con los datos presentados y discutidos se puede establecer que la utilización de técnicas de regresión junto al uso de Sistemas de Información Geográfica son una alternativa más que viable para solventar el problema de la actualización de los valores de inmuebles con fines contributivos. En varios de los modelos internacionales analizados estas técnicas son las utilizadas para actualizar los valores inmuebles con fines contributivos. Las pruebas realizadas demuestran un desempeño que cumple con los límites establecidos para las mismas.

Se puede observar cómo, sostenidamente por tres años en al menos 4 de los años bajo estudio, los valores de la prueba del Coeficiente de Dispersión cumplen con el límite máximo aceptable del COD. Los valores de venta del año 2007 no permitieron que esta condición se reflejara en los cinco años analizados, pero esto no debe sorprender porque los años seleccionados al incluir el año del estallido de la burbuja inmobiliaria se esperaba que este fuera el comportamiento de los resultados. Muchos territorios o países no siguen estrictamente la regla del 15% establecido por el IAAO y utilizan cifras que en ocasiones llega a un 20%. De elevar el valor adoptado del COD de un 15% a un 20%, los valores de predicción tendrían una utilidad cercana a los cinco años como base para el establecimiento de los valores sobre los cuales se establecería la tasa inmueble. De los modelos internacionales analizados se pudo observar la exigencia legal de actualización de valores catastrales gira entre los 3 a los 5 años. Adoptar un límite del Coeficiente de Dispersión de un 20% permitiría utilizar los datos de predicción al menos unos 5 años antes de realizar una nueva valorización.

Capítulo 8

Conclusiones

8. Conclusiones

La motivación principal que ha guiado esta investigación ha sido la situación en que se encuentra en la actualidad el modelo de valoración de inmuebles con fines contributivos en Puerto Rico. Los criterios que se utilizan para la valoración y cobro de contribuciones sobre los inmuebles no permiten que el sistema sea justo ni equitativo ya que estos datan de un procedimiento establecido en 1951 y ajustado en 1957, esto es, hace más de cincuenta años, sin revisión de los mismos. Además, dado que no se realiza un proceso de actualización de los datos y los dueños de inmuebles no informan de cambios estructurales en términos de mejoras como puede ser ampliaciones, construcción de terrazas o de piscinas, estos cambios estructurales del inmueble no están contemplados en el valor de tasación asignado. Se tributa basado en el valor asignado originalmente al inmueble una vez se inscribió y se valoró por primera vez con las tasas establecidas en 1957.

Esta inquietud llevo a establecer como uno de los objetivos de la presente investigación el seleccionar un método que permita establecer una base de valores de inmuebles residenciales con fines contributivos más precisa. El concepto adoptado para realizar la valoración se fundamentó en la valoración hedónica y la selección se realizó evaluando que técnica de regresión permite modelar mejor los datos seleccionados. Para esto se comparó la técnica de los Mínimos Cuadrados Ordinarios (MCO) y la técnica de Regresión Geográficamente Ponderada (RGP). Se utilizó Sistemas de Información Geográfica como la herramienta principal para ejecutar los modelos. Con esta herramienta se procesaron datos de ventas residenciales en un período de cinco años (2005 al 2009) y se aplicaron estas dos técnicas. De los resultados obtenidos se concluyó que la Regresión Geográficamente Ponderada permite realizar una predicción más certera de los datos. Estos resultados son cónsonos con los obtenidos por investigadores en trabajos similares (Brunsdon et al., 1999; Farber y Yeates, 2006; Bitter et al., 2007; Ogneva-Himmelberger et al., 2009; Sun y Jones, 2013; Lu et al., 2014; Dziauddin et al., 2015).

Quedo evidenciado en el estudio la importancia del factor localización en la aplicación de las técnicas de valoración hedónica. Al respecto Bitter et al. (2007), establecen que la localización de las ventas define la estructura espacial del mercado inmueble lo que ayuda a explicar las diferencias en valor. Esto ayuda a obtener coeficientes más certeros cuando se utiliza la técnica de valoración hedónica (Bitter et al., 2007). Sin embargo, uno de los problemas que trae esto son los efectos espaciales que presentan los mercados inmobiliarios,

en especial el de heterogeneidad espacial y autocorrelación espacial. Estos dos efectos, identificados en esta investigación, se convierten en los principales retos en el análisis de datos espaciales (Du y Mulley, 2006). Señala Bitter et al. (2007) que estos efectos surgen por el desbalance existente entre la localización de la oferta de los inmuebles y la demanda por estos. Con el objetivo de atender esta situación, una creciente cantidad de investigaciones (Brunsdon et al., 1999; Bitter et al., 2007; Yu, 2007; Ogneva-Himmelberger et al., 2009; Lu et al., 2014) han discutido la ventaja que presenta, al igual que la presente investigación, utilizar la técnica de la Regresión Geográficamente Ponderada. Esta técnica presenta una ventaja significativa cuando se compara con los Mínimos Cuadrados Ordinarios debido a la capacidad de estimar los parámetros según varían en el espacio. Esto provee una forma de ajustar las influencias de variables que se utilizan para la predicción en función del contexto espacial donde se ejercen las mismas. De las dos condiciones, la de la autocorrelación espacial presentó el mayor problema en esta investigación. El efecto de ésta se redujo limitando la cantidad de casos utilizados para realizar la predicción, pasando del universo de los casos, a utilizar solo una muestra de ellos. Otro problema como el de la redundancia de variables explicables se mantenía en el modelo. La reducción de variables explicativas eliminando aquellas que resultaban ser redundantes permitió corregir esta situación. Estas modificaciones al modelo inicial permitieron que la técnica de los MCO presentara un mejor desempeño. Sin embargo, al aplicar la técnica de RGP, los resultados presentaron un desempeño superior a los obtenidos con los MCO.

Unido a seleccionar la mejor técnica para modelar los datos, la investigación buscaba establecer la validez temporal de los valores de inmuebles residenciales como base para el componente contributivo del inmueble. Es claro que el proceso de valoración de inmuebles es uno costoso. Repetir el ejercicio de valoración anualmente representaría una fuerte inversión económica de parte del gobierno. La crisis económica por la que atraviesa el gobierno de Puerto Rico dificultaría el objetivo de realizar este ejercicio anualmente. Por tal razón se busca estudiar los resultados de las predicciones y probar con un método la validez de la utilización de estos por un período mayor al año. Este segundo objetivo de la investigación se manejó obteniendo datos de ventas residenciales para los cinco años posteriores a los seleccionados para el análisis de los modelos; esto es, datos de ventas residenciales para la zona de estudio entre los años 2010 y 2014. Se tomaron los datos de predicción de los años 2005 al 2009 y se asumió que estos serían las predicciones de valor inmueble de referencia para los años subsiguientes. Calculando el Coeficiente de Dispersión máximo esperado según

planteado por Gloudemans (2001) para cada año, se analizaron los resultados y se concluyó que las predicciones pueden mantener su validez por un período de tres años. Este es un período adecuado y que es igual al adoptado en varios estados en los Estados Unidos de América con es el estado de Texas (Commeaux, 2015) o que tan solo se diferencia por un año de los cuatro que adoptan Luisiana (Commeaux, 2015). Con un período de cuatro años también se observa el caso de Brasil (de Cesare, 2015).

En varias instancias de la investigación se discutió la influencia que eventos extremos que impactan el mercado inmobiliario pueden generar en el proceso de valoración (Bowman et al., 1986; Stevenson, 2008). El establecimiento del período de vigencia de las predicciones de tres años asumido en esta investigación podría ser cuestionado dado que los datos del año 2007 interrumpen el patrón de predicciones que cumplen con los estándares adoptados del COD. Se establece que el comportamiento irregular de los datos para el año 2007 se debe al estallido de la burbuja inmobiliaria que atravesaba el mercado inmobiliario de Puerto Rico. Es claro que el comportamiento irregular de ventas de inmuebles observado para este año afecta de alguna manera el cálculo del COD.

Este efecto en el valor del COD puede surgir por diversas situaciones. Por ejemplo, Levitt y Syverson (2008) concluyen que en mercados que no operan bajo condiciones normales, los agentes inmobiliarios cuentan con información privilegiada sobre el comportamiento del mercado que les brinda oportunidades al momento de realizar la transacción de venta de un inmueble. Esto presenta una ventaja cuando las gestiones inmobiliarias son ejecutadas por los propietarios de los inmuebles. En su estudio concluyen que inmuebles vendidos por agentes inmobiliarios en mercados que no operan bajo condiciones normales superan en valor de venta por un 3,7% a aquellas residencias vendidas directamente por sus propietarios. Por otro lado, las ejecuciones hipotecarias que surgen como consecuencia de la inestabilidad del mercado tienen efectos sobre los valores de mercado de otras propiedades en la jurisdicción (Frame, 2010). Estas condiciones se reflejan en los listados de ventas de inmuebles, los cuales se convierten en la fuente principal de datos al ejecutar el modelo hedónico.

Un efecto no relacionado directamente con los resultados del COD, pero discutido en la literatura surge con los cambios en los valores inmuebles que genera la situación económica. Las ventas inmuebles sirven de base para adoptar la tasa a la que se realiza el cobro de los valores inmuebles. Una adopción a largo plazo de la tasa provoca que el valor de ésta se

mantenga inalterado, aun cuando el costo de mercado de los inmuebles se reduzca (Anderson, 2010; Alm et al. 2014).

Es meritorio aclarar que el período de años seleccionado se eligió con la intención de identificar el impacto del estallido de la burbuja inmobiliaria en el modelo de valoración. La investigación partió con la idea que el efecto se manifestaría en el modelo, lo que no se conocía era su magnitud. Los resultados obtenidos no mostraron que el impacto fuera significativo ya que solo afectó uno de los años bajo análisis.

La experiencia de la investigación, los resultados obtenidos y los datos obtenidos al revisar modelos adoptados en otros países, ayudan a generar una serie de sugerencias que deben ser adoptadas localmente y que permiten acoger un modelo de valoración actualizado que viabilice el cobro de impuestos por propiedad inmueble de manera justa y real (cuarto objetivo).

Un aspecto no considerado dentro del modelo desarrollado en esta investigación es el de las zonas homogéneas (Montaña, 2016; Orrengo, 2016). Los países que utilizan este concepto buscan valores de predicción certeros amparados en la idea que la predicción se realiza solamente con datos muestreados para una zona homogénea. De adoptar este modelo en Puerto Rico se puede utilizar como región la zona de la urbanización ya que el desarrollo urbano en los pasados 70 años se ha dado utilizando este concepto, que de por sí, representa una zona económica, social y urbanísticamente homogénea. Aunque un concepto que cuenta con adeptos en el proceso de implantación, el mismo no está ausente de críticas (Goodman y Thibodeau, 1995; Yu, 2007). Al respecto Yu (2007) plantea que al aplicar técnicas de regresión, la agrupación espacial de datos da paso a lo que él llama modelación multinivel. El problema de éste es que la agrupación espacial de los datos presume conocimiento previo del mercado que se analiza y por medio de la agrupación ese conocimiento es introducido al modelo. Por lo tanto, se estarían introduciendo al modelo límites discretos que podrían acentuar el problema de la no-estacionariedad de los datos. Sería poco realista asumir límites discretos en los procesos espaciales de la dinámica del mercado inmobiliario, ya que el comportamiento de éste en el espacio es uno continuo (Fotheringham et al., 2002). La adopción tendría que asumir que cada zona homogénea es una célula totalmente aislada del resto del mercado de la ciudad, quedando fuera del análisis los efectos que busca modelar la Regresión Geográficamente Ponderada y que han sido discutidos. Por tal razón, antes de

adoptar el concepto de las zonas homogéneas, se sugiere realizar estudios más profundos que concluyan el beneficio que puede acarrear su adopción.

Otro aspecto observado en los modelos estudiados es el de la tasa aplicable al valor catastral. Se pudo observar a nivel internacional que cada país tiene una regla diferente. Sin embargo, se puede concluir que en todos los modelos estudiados esta tasa presentaba un valor máximo promedio en torno al 1,5% del valor catastral o tributable. Es importante aclarar que existe una diferencia marcada en todos los modelos estudiados entre el concepto de valor catastral y valor real. Se establece que el valor catastral siempre es menor al valor real. Incluso en los casos de Colombia (Montaña, 2016) y España (Palomo y Muller, 2010) específicamente se establece que los valores catastrales nunca superarán el valor real o el valor de mercado del inmueble. En el caso colombiano específicamente se establece que este valor puede llegar a ser hasta un 60% del valor de mercado del inmueble (Montaña, 2016). Con estas consideraciones el modelo a ser utilizado en Puerto Rico debe adoptar el concepto del valor catastral y establecer una diferencia clara con el concepto de valor real o valor de mercado. El posible error que surge del modelo de valoración podría muy bien quedar cubierto dentro de la especificación técnica de la diferencia entre ambos valores. En todo caso el estudio de los modelos de otros países lo que lleva a concluir es que el valor de mercado siempre es un punto de referencia para establecer el valor catastral y el valor por el cual se va a cobrar el impuesto al predio o inmueble. En el caso de España, el Texto Refundido de la Ley del Catastro Inmobiliario establece que el impuesto sobre los inmuebles se tributa basado en el valor catastral del inmueble. Este es un valor administrativo que se determina por la suma de los valores catastrales del suelo y de las edificaciones que hay sobre el mismo. Estos se calculan por un método de valoración que toma en consideración la localización del inmueble, parámetros urbanísticos, costes de edificación y situación de mercado. Pero más importante que esto es que el valor debe mantener una relación con el valor de mercado, no pudiendo superar el 50% de éste en el momento de aprobación de la ponencia. Por lo tanto, el valor de mercado sirve de referencia para el valor catastral del inmueble, no es igual ni le sustituye.

Común entre los modelos de valoración consultados se encuentran: la separación del valor suelo y el valor estructura, la utilización de sistema de clasificación de inmuebles diferenciando residenciales de comerciales o industriales y la aplicación de tasas en función del uso al que se dedica el inmueble. Todos los modelos consultados presentan procesos de

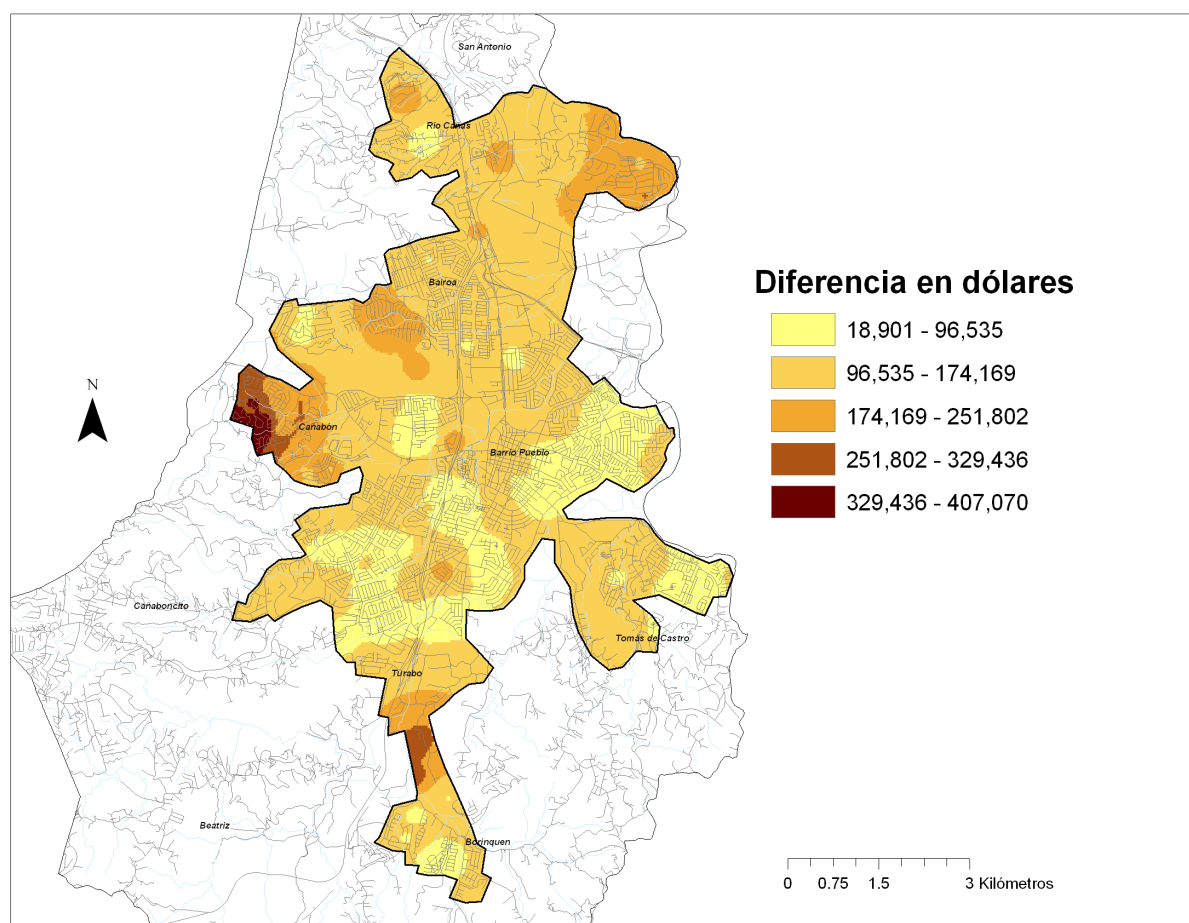
refutación. Con éste, el ciudadano que no está conforme con la valoración asignada a su inmueble puede solicitar se revise la misma.

A pesar de ser una medida antipática para los ciudadanos, los municipios en Puerto Rico adoptarían un proceso para la actualización de valores inmuebles con el objetivo de aumentar el recaudo de fondos en sus respectivos territorios. Para el año fiscal 2016-2017 se estima el ingreso del municipio de Caguas por concepto de impuesto a la propiedad inmueble en 32 millones de dólares. Se estima que para el año 2016, la diferencia entre el valor catastral y el valor de mercado en el municipio tiene una mediana de \$213.000 (Ver Mapa 8.1). Del total de viviendas en el municipio en el año 2016, unas 55.000 están sujetas al proceso de cobro de impuestos. La ley actual sobre valoración contempla una exoneración de 15.000 dólares a los hogares que son la residencia principal de la familia. Estas familias solo tributan por el excedente de los 15.000 dólares. Segundos hogares no reciben este beneficio. La Oficina del Censo de los Estados Unidos estima que, en Caguas en el año 2015, el porcentaje de residencias ocupadas por su dueño es el 68%. Se traduce este porcentaje en 37.400 residencias que están sujetas a la exoneración y 17.600 no. La cifra de exoneración responde al modelo con los valores basados en el costo de construcción del 1957. Para actualizar ésta se puede tomar de referencia lo establecido en la Ley 7 del 9 de marzo del 2009, mejor conocida como “Ley Especial Declarando Estado de Emergencia Fiscal y Estableciendo Plan Integral de Estabilización Fiscal para Salvar el Crédito de Puerto Rico”. En esta ley, para la contribución especial de tres años se adoptó una exoneración de 150.000 por hogar principal. Tomando estos datos se puede estimar la cantidad de dinero que el municipio podría recaudar con el modelo propuesto.

Luego de aplicada la exoneración a los hogares meritorios de éste beneficio, quedaría disponible un total de 6.105 millones de dólares para el cobro de impuestos. El gobierno de Puerto Rico ha argumentado que, con el modelo actual, en promedio en la isla, solo se captura el 0,38 del valor de mercado de las propiedades (Tellado, 2017). En el caso del municipio de Caguas, para capturar los 32 millones de dólares estimados por concepto de impuesto a la propiedad inmueble en el presente año fiscal, la tasa contributiva tendría que ser de un 0,53 por ciento. El gobierno argumenta que, de adoptar un modelo nuevo, se aspira que la tasa parta de un 0,65 por ciento. Con esta cifra, se podría estimar el recaudo en el municipio de unos 39 millones de dólares. Si la tasa se eleva al 0,75 podría recaudar unos 45 millones de dólares y con una tasa de 0,80 por ciento se podría recaudar unos 48 millones de dólares.

La pregunta de rigor es ¿Cómo impacta esto en el ciudadano? Para esto se toma como ejemplo un hogar promedio. El inmueble cuenta con una residencia cuya superficie de ocupación es de 141 metros cuadrados localizados en un predio de 387 metros cuadrados. El valor catastral es de 18.086 dólares. Como es hogar principal se le aplica la exoneración de 15.000 dólares lo que deja un total de 3.086 dólares tributables. Aplicando la tasa de 9,83 prevaleciente en el municipio, el inmueble paga 303,35 dólares en impuestos al año. El valor de mercado del inmueble es de 168.850 dólares. Con una tasa de 0,50 porciento aplicado al valor de mercado el pago de impuestos del inmueble se eleva a 995 dólares anuales y con una tasa de 0,65 se elevaría a 1.225 dólares anuales. Si el hogar no fuera la residencia principal de la familia tendría que tributar 1.777 dólares anuales con el modelo actual y con el modelo propuesto a una tasa del 0,50 tributaría 844 dólares. Si la tasa es 0,65 tendría que pagar 1.097 dólares anuales.

MAPA 8.1: Diferencia entre valor de mercado y valor catastral. Año 2016



El proceso de actualizar el modelo de valoración de inmuebles con fines contributivos ha sido dilatado por un gran período de tiempo principalmente por decisiones de los gobernantes del país. Anunciar al pueblo un proceso de revalorización de los inmuebles con el objetivo de obtener más fondos para el erario público no es una noticia agradable. En más de una ocasión se ha discutido la necesidad de realizar esta actualización y son los propios alcaldes los principales opositores a estas medidas, dado su posible costo político. A pesar de ello, es un ejercicio que, llevado de la manera correcta, puede tener beneficios para la sociedad y no representar un impacto significativo para los dueños de los inmuebles. A esto se debe sumar la particular situación económica por la que atraviesa el país. El estallido de la burbuja inmobiliaria en el año 2007 tan sólo fue el comienzo de una recesión económica que ha estado sufriendo Puerto Rico por los pasados 10 años. La migración de puertorriqueños a los Estados Unidos ha hecho que en estos pasados 10 años cerca de 1 millón de personas migren a los Estados Unidos, mayormente profesionales de la clase trabajadora, los que en gran medida son la base del modelo impositivo del país. Con una clase productiva en reducción constante, cada vez es menor la cantidad de impuestos que el gobierno puede recolectar. Con un escenario como éste es mucho más baja la probabilidad que los alcaldes, quienes son los principales promotores del proceso de cobro por impuestos de inmuebles, impulsen un proyecto de actualización de valores inmobiliarios. Esto ante los ojos del pueblo sería una nueva medida para extraer más dinero del ya afectado bolsillo del ciudadano. Sin embargo, una reforma contributiva amplia, que no tan solo tome el aspecto de los valores de los impuestos por inmuebles, sino que también tome en consideración el de los impuestos al ingreso y al consumo y a la vez cree un balance entre estos tres principales impuestos, muy bien podría dar espacio para que se haga una revisión total del modelo de valoración por inmuebles y se reajuste el pago de impuestos por esta condición.

Al final, se pudo establecer un modelo que resuelve el problema de contar con una valoración inmueble actualizada de manera rápida cuando es comparado con el posible ejercicio de visitar cada inmueble para tasarlo. El modelo depende de la acción del mismo mercado inmobiliario, porque depende del listado de ventas inmuebles y sus características realizadas en meses anteriores. El modelo puede ser replicado no solo en otros municipios de Puerto Rico, sino también en otras regiones geográficas.

Bibliografía

- Acosta, I. (1995). *Santa Juana y Mano Manca: Auge y decadencia del azúcar en el Valle del Turabo en el siglo XX*. San Juan: Editorial Cultural.
- Alm, J., Buschman, R. D., & Sjoquist, D. L. (2014). How Do Foreclosures Affect Property Values and Property Taxes? *LandLines*. January 2014. Page 22
- Anderson, J. E. (2010). Shocks to the property tax base and implications for local public finance. *Paper presented at the Urban Institute-Brookings Institution Tax Policy Center and the Lincoln Institute of Land Policy Conference, "Effects of the Housing Crisis on State and Local Governments," Washington, D.C. (May)*.
- Anselin, L., & Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs*, 155, 237-290.
- Anselin, L., & Le Gallo, J. (2006). Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects. *Spatial Economic Analysis*, 1(1), 31-52.
- Anselin, L., & Lozano-Gracia, N. (2008). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, 34(1), 5-34.
- Appraisal Institute (2013). *The Appraisal of Real Estate* (14th edition ed.). Chicago: Appraisal Institute.
- Aravena, N. (2007). *Enfoques de Tasación*. Notas de curso: Tasación Inmobiliaria. Escuela de Arquitectura. Universidad de las Américas, Chile.
- Ballesteros, E., & Rodríguez, J. A. (1999). *El Precio de los Inmuebles Urbanos* (Segunda Edición ed.). Madrid: Cie Inversiones Editoriales Dossat.
- Base, B. (1991). The Dynamics of Real Estate Prices. *Review of Economics and Statistics*, 50-58.
- Basu, S., & Thibodeau, T. G. (1998). Analysis of Spatial Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, 17(1), 61-85.
- Benegas Lynch, A., & Dania, R. (2001). Sistemas Tributarios: un análisis en torno al caso argentino *Reformas Tributarias en América Latina, Buenos Aires*. Buenos Aires.
- Benmamoun, M. (2006). Bootstrap confidence intervals and Gloudehman's COD tolerance test using SPSS and Stata. *Journal of property tax assessment & administration*, 3(4), 51-51.
- Berenguer, C. D. (2009). La información catastral en la Tasación Inmobiliaria. *CT Catastro* (67), 137-143.
- Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, 9(1), 7-27.
- Bourassa, S., Cantoni, E., & Hoesli, M. (2007). Spatial Dependence, Housing Submarkets, and House Price Prediction. *The Journal of Real Estate Finance and Economics*, 35(2), 143-160.
- Bowman, J. H., & Butcher, W. A. (1986). Institutional remedies and the uniform assessment of property: An update and extension. *National Tax Journal*, 157-169.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4), 281-298.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1998). Geographically Weighted Regression-Modelling Spatial Non-Stationarity. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47(3), 431-443.

- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39(3), 497-524.
- Bunker, O. L. (1975). *Historia de Caguas*. España.
- Bunker, O. L. (1983). *Caguas: Notas para su Historia*. San Juan: Municipio de Caguas.
- Burt, J. E., Barber, G. M., & Rigby, D. L. (2009). *Elementary Statistics for Geographers*. New York: Guilford.
- Can, A. (1998). GIS and Spatial Analysis of Housing and Mortgage Markets. *Journal of Housing Research*, 9(9), 61-86.
- Caridad, J. M., Núñez, J. M., & Ceular, N. (2008). Metodología de precios hedónicos vs. Redes Neuronales Artificiales como alternativas a la valoración de inmuebles. Un caso real. *CT Catastro*, 62, 27-42.
- Código de Puerto Rico: Ley para establecer un código político para Puerto Rico. (1902) Puerto Rico.
- Comeaux, C. T. (2015). *Louisiana Property Tax Basics*. Lafayette: Lafayette Parish.
- Coy, P. (2008). Housing Meltdown. (Cover story). *Business Week* (4070), 040-046.
- De Cesare, C. M. (2016). Brasil. In C. M. d. Cesare (Ed.), *Sistemas del Impuesto Predial en América Latina y el Caribe*. Boston: Lincoln Institute of Land Policy.
- Devaux, N., & Dubé, J. (2016). About the Influence of Time on Spatial Dependence: A Meta-analysis using Real Estate Hedonic Pricing Models. *Journal of Real Estate Literature*, 24(1), 31-66.
- Dorsey, R. E., Hu, H., Mayer, W. J., & Wang, H.-c. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, 19(2), 75-93.
- Du, H., & Mulley, C. (2006). Relationship Between Transport Accessibility and Land Value: Local Model Approach with Geographically Weighted Regression. *Transportation Research Record: Journal of the Transportation Research Board*, 1977(-1), 197-205.
- Dubé, J., & Legros, D. (2014). Spatial econometrics and the hedonic pricing model: what about the temporal dimension? *Journal of Property Research*, 31(4), 333-359.
- Dubé, J., Thériault, M., & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992-2009. *Transportation Research Part A: Policy and Practice*, 54, 49-66.
- Dziauddin, M. F., Powe, N., & Albanides, S. (2015). Estimating the effects of light rail transit (LRT) system on residential property values using geographically weighted regression (GWR). *Applied Spatial Analysis and Policy*, 8(1), 1-25.
- Farber, S., & Yeates, M. (2006). A comparison of localized regression models in a hedonic house price context. *Canadian Journal of Regional Science*, 29(3), 405-420.
- Feng, X., & Humphreys, B. (2016). Assessing the Economic Impact of Sports Facilities on Residential Property Values A Spatial Hedonic Approach. *Journal of Sports Economics*, 1527002515622318.
- Ferrando-Correl, J. V. (2004). *Valoración de Inmuebles de Naturaleza Urbana*. Valencia: Servicio de Publicación de la Universidad de Valencia.
- Figueroa, E., & Lever, G. (1992). Determinantes del precio de la vivienda en Santiago: Una estimación hedónica. *Estudios de Economía*, 19(1), 67-84.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2002). *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley.
- Frame, W. S. (2010). Estimating the effect of mortgage foreclosures on nearby property values: A critical review of the literature. *Economic Review-Federal Reserve Bank of Atlanta*, 95(3), II.

- Gillard, Q. (1981). The effect of environmental amenities on house values: the example of a view lot. *The Professional Geographer*, 33(2), 216-220.
- Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, 23(1), 5-30.
- Gloudemans, R. J. (1999). *Mass Appraisal of Real Property*. Chicago: International Association of Assessing Officers.
- Gloudemans, R. J. (2001). Confidence intervals for the coefficient of dispersion: Limitations and solutions. *Assessment Journal*, 8(6), 23-27.
- Goodman, A. C., & Thibodeau, T. G. (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, 6(1), 25.
- Harris, B. H., & Moore, B. D. (2013). *Residential Property Taxes in the United States*. Retrieved from Urban-Brookings Tax Policy Center
- Hong, Y.-H., & Brubaker, D. (2010). Integrating the Proposed Property Tax with the Public Leasehold System. In J. Y. Man & Y.-H. Hong (Eds.), *China's Local Public Finance in Transition*. Cambridge: Lincoln Institute of Land Policy.
- Iregui, A. M., Melo, L., & Ramos, J. (2004). *El Impuesto Predial en Colombia: Evolución Reciente, Comportamiento de las Tarifas y Potencial de Recaudo*. Colombia: Banco de la República.
- Jacobus, C. (2010). *Real Estate Principles* (11 ed.). Ohio: Cengage Learning.
- Junta de Planificación de Puerto Rico (2014). *Estadísticas Seleccionadas sobre la Industria de la Construcción*. San Juan: Junta de Planificación.
- Kauko, T. (2003). On current neural network applications involving spatial modelling of property prices. *Journal of Housing and the Built Environment*, 18(2), 23.
- Kauko, T. (2014). How to Understand Residential Value and Valuation. *Critical Housing Analysis*, 1(1), 8.
- Kennedy, P. E. (2005). Oh no! I got the wrong sign! What should I do? *The Journal of Economic Education*, 36(1), 77-92.
- Kestens, Y., Thériault, M., & Des Rosiers, F. (2006). Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems*, 8(1), 61-96.
- Klovers, K., & Pereira, A. (2015). *Estimating the effect of transit on residential property values: the case of the Portland MAX system*. Retrieved from
- Krugman, P. (2009, 2 de Septiembre). How Did Economists Get It So Wrong? *New York Times*. Retrieved from <http://www.nytimes.com/2009/09/06/magazine/06Economic-t.html>
- Krugman, P. (2014, December 28). The Obama Recovery. *New York Times*. Retrieved from http://www.nytimes.com/2014/12/29/opinion/paul-krugman-the-obama-recovery.html?_r=0
- Kutasi, D., & Badics, M. C. (2016). Valuation methods for the housing market: Evidence from Budapest. *Acta Oeconomica*, 66(3), 527-546. doi:10.1556/032.2016.66.3.8
- Larson, W. (2015). *New Estimates of Value of Land of the United States*. Retrieved from Bureau of Economic Analysis: <http://www.bea.gov/papers/pdf/new-estimates-of-value-of-land-of-the-united-states-larson.pdf>
- Larsson, G. (1991). *Land Registration and Cadastral Systems: Tools for Land Information and Management*. Harlow (Essex): Longman Scientific and Technical.
- Levitt, S. D., & Syverson, C. (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *The Review of Economics and Statistics*, 90(4), 599-611.
- Ley 117: Ley de Contribución sobre la Propiedad (1947). Gobierno de Puerto Rico
- Ley de Contribución Municipal sobre la Propiedad (1991). Gobierno de Puerto Rico

- Ley Especial Declarando Estado de Emergencia Fiscal y Estableciendo Plan Integral de Estabilización Fiscal para Salvar el Crédito de Puerto Rico (2009). Gobierno de Puerto Rico
- Li, M. M., & Brown, H. J. (1980). Micro-neighborhood externalities and hedonic housing prices. *Land economics*, 56(2), 125-141.
- Li, R. Y. M., Chau, K. W., Li, M. H., Leung, T. H., & Kwok, K. W. (2016). Negative environmental externalities and housing price: A hedonic model approach *Econometric Analyses of International Housing Markets* (pp. 101-115).
- Liu, Y., Zheng, B., Turkstra, J., & Huang, L. (2010). A hedonic model comparison for residential land value analysis. *International Journal of Applied Earth Observation and Geoinformation*, 12(Supplement 2), S181-S193.
- Lu, B., Charlton, M., Harris, P., & Fotheringham, A. S. (2014). Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science*, 28(4), 660-681.
- Marxuach, S. M. (2010). El desfase del mercado de vivienda. Retrieved from <http://grupocne.org/2010/08/13/el-desfase-del-mercado-de-vivienda/>
- McCluskey, W. J. (2005). *Land value Taxation: An applied analysis / edited by William J. McCluskey, Riël C.D. Franzsen*. Aldershot : Ashgate.
- Miranda, B. B. (2012). Impuesto sobre bienes inmuebles y valor catastral: una reforma pendiente. *CT Catastro* (75), 7-30.
- Molinatti, C. (2016). Argentina. In C. M. d. Cesare (Ed.), *Sistemas del Impuesto Predial en América Latina y el Caribe*. Boston: Lincoln Institute of Land Policy.
- Montaña, M. (2016). Colombia. In C. M. d. Cesare (Ed.), *Sistemas del Impuesto Predial en América Latina y el Caribe*. Boston: Lincoln Institute of Land Policy.
- Morano, P., & Tajani, F. (2014). Least median of squares regression and minimum volume ellipsoid estimator for outliers detection in housing appraisal. *International Journal of business intelligence and data mining*, 9(2), 91-111.
- Mulley, C., & Tsai, C. H. (2017). Impact of bus rapid transit on housing price and accessibility changes in Sydney: A repeat sales approach. *International Journal of Sustainable Transportation*, 11(1), 3-10. doi:10.1080/15568318.2015.1106223
- Oficina del Contralor de Puerto Rico (2014). *Informe de Auditoría al Centro de Recaudación de Ingresos Municipales*. (TI-14-21). Puerto Rico Retrieved from http://www.ocpr.gov.pr/informes_en_PDF/pdf_2013_2014/ti/TI-14-21.pdf.
- Ogneva-Himmelberger, Y., Pearsall, H., & Rakshit, R. (2009). Concrete evidence & geographically weighted regression: a regional analysis of wealth and the land cover in Massachusetts. *Applied Geography*, 29(4), 478-487.
- Orrego, C. O. (2007). *El Impuesto Territorial en Chile*. Retrieved from Lincoln Institute of Land Policy - Boston:
- Orrego, C. O. (2016). Chile. In C. M. d. Cesare (Ed.), *Sistemas del Impuesto Predial en América Latina y el Caribe*. Boston: Lincoln Institute of Land Policy.
- Orford, S. (1999). *Valuing the Built Environment: GIS and house price analysis*. Aldershot: Ashgate.
- Palomo, M. D. C., & Müller, A. (2010). *Comparación entre la valoración catastral de los bienes inmuebles en España y en Alemania: El valor catastral español frente al valor unitario (Einheitswert) alemán*. Paper presented at the Catastro: formación, investigación y empresa: Selección de ponencias del I Congreso Internacional sobre catastro unificado y multipropósito
- Peterson, S., & Flanagan, A. B. (2009). Neural network hedonic pricing models in mass real estate appraisal. *Journal of Real Estate Research*, 31(2), 147-164. Retrieved from

- <https://www.scopus.com/inward/record.uri?eid=2-s2.0-70350502033&partnerID=40&md5=aff4bac9ea2ec044c51623c29c599326>
- Plan de Ordenación Territorial de Caguas (2011). Municipio Autónomo de Caguas. Puerto Rico.
- Quercia, R. G., McCarthy, G. W., Ryznar, R., & Can, A. (2000). Spatio-Temporal Measurement of House Price Appreciation in Underserved Areas. *Journal of Housing Research*, 11(1), 943-972.
- Rattermann, M. (2004). *The Student Handbook to The Appraisal of Real Estate*. Chicago: Appraisal Institute.
- Reglamento Para Establecer Guías y Procedimientos En La Valoración De La Propiedad Inmueble Por Su Destino Para Fines Contributivos (2008). Centro de Recaudación de Ingresos Municipales.
- Ridker, R. G., & Henning, J. A. (1967). The Determinants of Residential Property Values with Special Reference to Air Pollution. *The Review of Economics and Statistics*, 49(2), 246-257. Retrieved from <http://www.jstor.org/stable/1928231>
- Rodriguez, M., & Sirmans, C. (1994). Quantifying the value of a view in single-family housing markets. *Appraisal Journal*, 62, 600-600.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34-55.
- Rosenshein, L., Scott, L., & Pratt, M. (2011). Finding a Meaningful Model. *ArcUser*, Winter 2011, 5.
- Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. *Journal of Transport Geography*, 41, 53-62.
- Shiller, R. J. (2011, Febrero 5). Housing Bubbles Are Few and Far Between. *New York Times*.
- Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature*, 13(1), 1-44.
- Sirpal, R. (1994). Empirical modeling of the relative impacts of various sizes of shopping centers on the values of surrounding residential properties. *Journal of Real Estate Research*, 9(4), 487-505.
- Smith, M. d., Longley, P., & Goodchild, M. (2015). *Geospatial Analysis - A Comprehensive Guide to Principles, Techniques and Software Tools*. UK: The Winchelsea Press.
- Stevenson, S. (2008). Modeling Housing Market Fundamentals: Empirical Evidence of Extreme Market Conditions. *Real Estate Economics*, 36(1), 1-29.
- Sun, H., Wang, Y., & Li, Q. (2016). The Impact of Subway Lines on Residential Property Values in Tianjin: An Empirical Study Based on Hedonic Pricing Model. *Discrete Dynamics in Nature and Society*, 2016.
- Sun, W., & Jones, B. (2013). Using multi-scale spatial and statistical analysis to assess the effects of brownfield redevelopment on surrounding residential property values in Milwaukee County, USA. *Moravian Geographical Reports*, 21(2), 56-64.
- Suriatini, I. (2006). Spatial Autocorrelation and Real Estate Studies: A Literature Review. *Malaysian Journal of Real Estate*, 1(1), 1- 13.
- Swoboda, A., Nega, T., & Timm, M. (2015). Hedonic analysis over time and space: the case of house prices and traffic noise. *Journal of Regional Science*, 55(4), 644-670.
- Tellado, R. N. (2017, 3/5/2017). El mercado hipotecario enfrenta una encrucijada. *El Nuevo Día*.
- Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, 234-240.

- Tu, Y., Sun, H., & Yu, S.-M. (2007). Spatial Autocorrelations and Urban Housing Market Segmentation. *The Journal of Real Estate Finance and Economics*, 34(3), 385-406.
- Tu, Y., Yu, S.-M., & Sun, H. (2004). Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach. *Real Estate Economics*, 32(2), 297-328.
- Ventolo, W. L., & Williams, M. R. (2012). *Fundamentals of Real Estate Appraisal* (11 ed.). Wisconsin: Dearborn Real Estate Education.
- Walden, M. (1990). Magnet schools and the differential impact of school quality on residential property values. *Journal of Real Estate Research*, 5, 221-230.
- Wee, S. (2016). The effect of residential solar photovoltaic systems on home value: A case study of Hawai'i. *Renewable Energy*, 91, 282-292.
- Yaolin, L., Tao, D., Yanfang, L. (2006). *A GIS-based hedonic modeling of urban land value spatio-temporal patterns* (Vol. 6366). Bellingham, WA, ETATS-UNIS: Society of Photo-Optical Instrumentation Engineers.
- Yu, D. (2007). Modeling owner-occupied single-family house values in the city of Milwaukee: A geographically weighted regression approach. *GIScience & Remote Sensing*, 44(3), 267-282.
- Zhong, H., & Li, W. (2016). Rail transit investment and property values: An old tale retold. *Transport Policy*, 51, 33-48.
- Zhou, P., Liu, Y., Chen, Y., Zeng, C., & Wang, Z. (2015). Prediction of the spatial distribution of high-rise residential buildings by the use of a geographic field based autologistic regression model. *Journal of Housing and the Built Environment*, 30(3), 487-508.
- Zygmunt, R., & Gluszak, M. (2015). Forest proximity impact on undeveloped land values: A spatial hedonic study. *Forest Policy and Economics*, 50, 82-89.

Anexo

ANALISIS DE LA INFLUENCIA DE COMPONENTES GEOGRAFICOS Y TEMPORALES EN VALORES RESIDENCIALES UTILIZANDO SIG

Semana Geomática Internacional 2013, Instituto Geográfico Agustín Codazzi (IGAC). Bogotá, Colombia

Raúl MATOS FLORES

Depto. de Ciencias Geomáticas. Universidad Politécnica de Puerto Rico

ramatos@pupr.edu

RESUMEN

La valoración hedónica tiene como objetivo establecer el valor de una variable dependiente basándose en una serie de características del mismo (variables independientes o explicativas). La identificación de éstas variables independientes o explicativas va en función de la hipótesis que se formula sobre la influencia de los diversos factores que conforman el valor buscado. Se busca con esta investigación identificar la combinación ideal de variables con la que se puede realizar una predicción de valor residencial para la zona de estudio.

Palabras clave: Mínimos Cuadrados Ordinarios, Sistema de Información Geográfica, Valor Hedónico, Valor Residencial

ABSTRACT

Hedonic valuation aims to establish the value of a dependent variable based on a series of characteristics that variable. The identification of these explanatory variables depends on the formulated hypothesis on how those factors influence on the value sought. This research looks to identify the ideal combination of variables with which a residential value prediction can be made.

Key words: Geographic Information Systems, Hedonic Valuation, Land Value, Ordinary Least Squares.

1. INTRODUCCION

La valoración hedónica tiene como objetivo establecer el valor de una variable dependiente basándose en una serie de características del mismo (variables independientes o explicativas). La identificación de éstas variables independientes o explicativas va en función de la hipótesis que se formula sobre la influencia de los diversos factores que conforman el valor buscado. En ocasiones la selección y utilización de estas variables depende del acceso a los datos. Sobre la selección de éstas, Caridad et al. (2008) plantean que es deseable incluir un

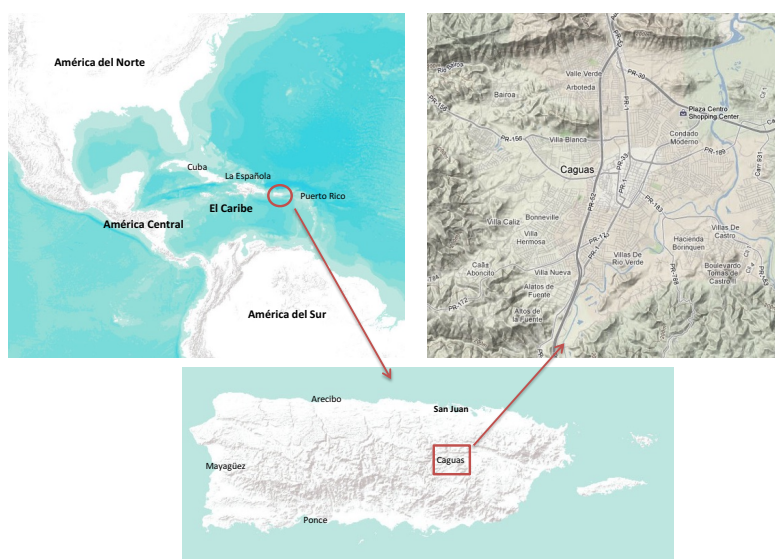
número no demasiado elevado de este tipo de variables por dos razones principales; primero, porque las variables independientes suelen estar relacionadas entre sí y, en segundo lugar, porque la inclusión de un número elevado de variables no origina importantes mejoras en el poder explicativo del modelo.

De todas las técnicas de regresión utilizadas para el modelo de valoración hedónica la más conocida es la de los Mínimos Cuadrados Ordinarios (Yaolin, 2006; Kestens et al., 2006; Liu et al., 2010; Dorsey et al., 2010; Seo et al., 2014; Morano y Tajani, 2014; Klovers y Pereira, 2015). La técnica de los Mínimos Cuadrados Ordinarios (MCO) se considera un punto de inicio adecuado para todos los análisis de regresión espacial. Esta técnica proporciona un modelo global de la variable que intenta predecir ya que crea una ecuación de regresión simple para representar ese proceso.

2. ZONA DE ESTUDIO

Para este estudio se seleccionó el municipio de Caguas, Puerto Rico (Figura 2.1). El municipio de Caguas está localizado en la porción central este de Puerto Rico, contando con una extensión aproximada de 152 km² (un 1,7% del territorio de Puerto Rico). Caguas se localiza al sur de San Juan (Capital de Puerto Rico) y su Región Metropolitana. Sus características fisiográficas principales son: las tierras planas y semiplanos del valle central, las colinas al norte del valle y la zona más escarpada hacia el sur y el oeste. El valle ocupa aproximadamente 41 km² y se encuentra a una elevación promedio de 60 m.s.n.m.. Las colinas al norte del valle cuentan con elevaciones entre los 90 y 240 m. y las zonas más escarpadas al sur y oeste con elevaciones máximas de 600 m.s.n.m.

Figura 2.1: Localización municipio de Caguas, Puerto Rico



Según el Censo de Población y Vivienda de 2010, Caguas cuenta con 142.893 residentes y un total de 60.371 unidades de vivienda.

3. SELECCIÓN DE VARIABLES

La aplicación de un modelo hedónico obliga a la identificación de una serie de atributos importantes para la predicción del precio del inmueble. Esto crea la necesidad de recopilar datos que permitan ejecutar el modelo y obtener la predicción de valor buscado. En el caso específico de los valores de inmuebles, la especificación hedónica asume que los atributos seleccionados representan características físicas del inmueble, características relacionadas al vecindario y características relacionadas a la localización del inmueble.

Los datos seleccionados corresponden al período entre los años 2005 y 2009. Se seleccionó este período de manera que se pudiera modelar la situación del mercado de ventas de inmuebles residenciales por el que atravesó Puerto Rico, en particular reflejar el estallido en el año 2007 de la burbuja inmobiliaria por la que atravesaba la isla (Shiller, 2007; Krugman, 2009). Se busca identificar la influencia de esta situación en el modelo.

Los datos se obtuvieron de listados producidos por firmas especializadas en generar listados de ventas de inmuebles que mayormente son utilizadas por tasadores en sus funciones de valoración. Los datos para producir estos listados surgen de los informes mensuales producidos por las entidades bancarias que generaron los préstamos hipotecarios. Estos listados de datos no solo incluyen una serie de variables que fueron consideradas por el tasador al momento de generar una opinión de valor sobre el inmueble, sino que también incluye datos tan variados como el nombre del tasador que generó la tasación, el estimado de valor, el valor de la venta final, el banco que generó la hipoteca y otros.

De los datos obtenidos de la base de datos se seleccionaron las más representativas de las características físicas del inmueble, otras que presentan características de la comunidad donde se localizan los inmuebles y se creó una tercera donde se calculó la distancia de los inmuebles a diversos servicios. Las variables seleccionadas al final fueron:

- Variable dependiente
 - Saleprice: Precio de venta del inmueble en dólares.
- Variables Independientes
 - Estructurales de la residencia
 - Lotarea: Superficie del solar medido en metros cuadrados.
 - Age: Años que lleva construido el inmueble.
 - Bedrooms: Cantidad de dormitorios con que cuenta el inmueble.
 - Bathrooms: Cantidad de baños con que cuenta el inmueble.
 - HouseareaM: Superficie de ocupación del inmueble medido en metros cuadrados.
 - Comunitarios
 - ACC: Si la comunidad es una cerrada y cuenta con control de acceso.
 - GZB: Si las facilidades recreativas de la comunidad cuentan con un gazebo¹⁰ para actividades.

¹⁰ Un gazebo es un pabellón de planta simétrica, generalmente hexagonal o circular, que comúnmente se encuentra en los parques, jardines, y en áreas públicas abiertas. Los gazebos se encuentran aislados, techados, y

- PG: Existencia de área de juego para niños.
- SWP: Existencia de piscina.
- TENIS: Existencia de canchas de tenis.
- BSKT: Existencia de canchas de baloncesto.
- PP: Existencia de un parque pasivo.
- GYM: Existencia de gimnasio.
- CH: Si el complejo residencial tiene una Casa Club para reuniones.
- VB: Existencia de cancha de voleibol.
- BB: Existencia de parque de béisbol.
- GLF: Existencia de campo de golf.
- Distancia a servicios (Localización)
 - CC_DIST: Distancia euclidiana a centro comercial más cercano.
 - EXIT_DIST: Distancia euclidiana a conector de vía de circulación nacional más cercano.

4. DIAGNOSTICO

Estas variables fueron sometidas a un proceso de diagnóstico para determinar finalmente cuales serían utilizadas en el modelo hedónico. Previo a este diagnóstico los valores de ventas fueron deflactados utilizando los datos del Índice de Precios al Consumidor publicados por el Departamento del Trabajo de Puerto Rico. Para el diagnóstico se realizaron una serie de pruebas estadísticas que ayudaron a determinar el desempeño general del modelo. Las pruebas fueron:

- Coeficiente de determinación múltiple R^2 y la del R^2 ajustado.
- Factor de Inflación de la Varianza (FIV) para medir la redundancia entre las variables explicativas
- Prueba Jarque – Bera para medir la distribución de los valores residuales.
- Índice de Moran

Las pruebas diagnósticas realizadas muestran varios problemas, principalmente uno de alta correlación espacial en las ventas. Ya se planteó anteriormente que los residuales producidos por un modelo de valoración hedónica aplicado a inmuebles pueden estar espacialmente correlacionados por tres razones (Gillen et al., 2001; Tu et al., 2004; Suriatini, 2007). Primero por las características estructurales del inmueble y las características del vecindario donde se localiza, segundo por el proceso de valoración utilizado para determinar el precio del inmueble y tercero por malas especificaciones del modelo MCO.

Basado en estos resultados se puede asumir que existe un problema de predictores redundantes (Rosenshein et al., 2011). Esa redundancia puede dar paso a la determinación de autocorrelación espacial por las respectivas pruebas. En los datos utilizados para el análisis se registran ventas en urbanizaciones nuevas. Dependiendo del tipo de inmueble en un período de venta de un año se puede entregar entre 20 y 100 unidades de vivienda en un nuevo

abiertos por todos los lados; proporcionan sombra, un abrigo de características básicas, función ornamental en un paisaje, y un lugar de descanso.

complejo residencial. En ocasiones esta cifra puede ser superior. Dada las características de los inmuebles, los casos utilizados para el análisis muestran poca o ninguna variación en sus características estructurales. Por otro lado, se observa una alta multicolinealidad en algunas variables del modelo. La multicolinealidad es una característica de la muestra seleccionada y no de los errores de estimación. Uno de los factores que impulsa a esta condición cuando alguna o todas las variables explicativas están altamente correlacionadas entre sí.

La utilización de la totalidad de ventas de inmuebles para cada año de estudio puede provocar estos resultados irregulares. Para corregir esta situación se decidió utilizar un método donde se trabajará con una muestra de los datos y no el universo. Para esto se realizaron dos muestras de los datos. La primera fue obtenida por medio de un muestreo aleatorio y el segundo una representación de los valores medianos de venta y características de cada comunidad. Con esto se busca eliminar la redundancia de casos que afectan los resultados esperados. La muestra aleatoria se logró con la herramienta de muestreo que contiene las herramientas de trabajo Hawth's Tools¹¹. Ésta realiza una selección aleatoria de casos de una base de datos digital cuya representación espacial se da por medio de puntos. La cantidad de casos seleccionada corresponde a un porcentaje del total establecido por el usuario. En el caso de estudio se estableció que la muestra fuera un 30% del total de casos. La segunda muestra se constituyó de la selección de ventas que representaran el valor mediano para cada una de las comunidades identificadas en la zona de estudio. Dada la alta similitud en las características de los inmuebles en cada comunidad se asume que el inmueble cuyos valores se consideren como la mediana de todos los valores de venta es representativo de la comunidad.

Con estas dos selecciones de datos se procedió a realizar una comparación para determinar cuál mostraba el mejor desempeño. La comparación se fundamentó en tres factores. Primero el valor R^2 , segundo el nivel de autocorrelación espacial y tercero el Factor de Inflación de Varianza. Bajo las tres condiciones el mejor desempeño en los cinco años analizados se observó en los valores que representaban el valor mediano de las ventas por cada comunidad. Dado estos resultados se procedió a realizar el análisis exploratorio de datos, esta vez con la muestra seleccionada.

Con el objetivo de depurar a un grado mayor e identificar la combinación idónea de variables a utilizar en el modelo de predicción se aplicó a los datos el proceso de “Regresión Exploratoria” que contiene el programa ArcGIS. Esta herramienta evalúa todas las posibles combinaciones de predicción para una variable dependiente utilizando las variables independientes seleccionadas. Para ejecutar esta función se establecen los valores aceptados para cuatro pruebas exploratorias de datos que se realizan utilizando diversas combinaciones de variables independientes. Se busca que el modelo seleccionado cumpla con las siguientes condiciones:

- Que exceda un límite de R^2 Ajustado especificado
- Que los valores P de los coeficientes sean menores a los especificados
- Que los valores de coeficiente de FIV, para todas las variables explicativas, sea menor que el especificado
- Que cuente con un valor P Jarque-Bera más grande que el especificado

Para ejecutar la herramienta se seleccionaron las variables bajo análisis y se indicó que los parámetros a cumplir serían los siguientes:

¹¹ Se puede descargar en www.spatial ecology.com

- R^2 Ajustado: Mayor a 0,70
- FIV: Menor a 7,5
- P-valor: máximo de 0,05
- Valor P Jarque Bera: Mínimo aceptado 0,1
- Valor P Autocorrelación espacial: Mínimo aceptado 0,1

Una vez ejecutado el análisis exploratorio, se observó la selección de nueve variables de manera consistente en modelos para los diversos años. Tomando estas nueve variables se ejecutó nuevamente la técnica de los Mínimos Cuadrados Ordinarios para cada año, arrojando valores aceptables para los cinco años de estudio (Tabla 4.1). Se estableció así que éstas serían las variables finales para la comparación de técnicas. Las variables son:

- Estructurales de la residencia
 - LOTAREA: Superficie del solar medido en metros cuadrados.
 - AGE: Años que lleva construido el inmueble.
 - HOUSEAREAM: Superficie de ocupación del inmueble medido en metros cuadrados.
- Comunitarios
 - TENIS: Existencia de canchas de tenis
 - GYM: Existencia de facilidades de gimnasio para ejercicios
 - VB: Existencia de cancha de voleibol.
 - BB: Existencia de parque de béisbol.
 - GLF: Existencia de campo de golf.
- Distancia a servicios (Localización)
 - CC_Dist: Distancia euclidiana a centro comercial más cercano

TABLA 4.1: Resultados Modelo con Nueve (9) Variables Finales

		2005	2006	2007	2008	2009
R2		0,91	0,88	0,90	0,85	0,86
VIF (Mas Alto)		2,86	2,53	2,29	2,56	3,22
AiCc		2517,24	3131,70	2122,16	3799,74	1875,24
Koenker valor)	(P	0,0540	0,000047*	0,000934*	0,0000*	0,00005*
Jarque Bera Valor)	(P	0,8134	0,0000*	0,0000*	0,0000*	0,08043
Moran I		0,0711	-0,3292	-0,0224	0,0597	0,0408
Valor Z		1,039139	-0,546229	-0,15598	0,741716	0,764689
Valor P		0,29874	0,58490	0,87604	0,45825	0,44457

5. CONCLUSIONES

El estudio demostró la importancia del factor localización en la aplicación de las técnicas de valoración hedónica. Al respecto Bitter et al. (2007), establecen que la localización de las

ventas define la estructura espacial del mercado inmueble lo que ayuda a explicar las diferencias en valor. Esto ayuda a obtener coeficientes más certeros cuando se utiliza la técnica de valoración hedónica (Bitter et al., 2007). Sin embargo, uno de los problemas que trae esto son los efectos espaciales que presentan los mercados inmobiliarios, en especial el de heterogeneidad espacial y autocorrelación espacial. Estos dos efectos, identificados en esta investigación, se convierten en los principales retos en el análisis de datos espaciales (Du y Mulley, 2006). Señala Bitter et al. (2007) que estos efectos surgen por el desbalance existente entre la localización de la oferta de los inmuebles y la demanda por estos. Con el objetivo de atender esta situación, una creciente cantidad de investigaciones (Brunsdon et al., 1999; Bitter et al., 2007; Yu, 2007; Ogneva-Himmelberger et al., 2009; Lu et al., 2014) han discutido la ventaja que presenta utilizar la técnica de la Regresión Geográficamente Ponderada. Esta técnica presenta una ventaja significativa cuando se compara con los Mínimos Cuadrados Ordinarios debido a la capacidad de estimar los parámetros según varían en el espacio. Esto provee una forma de ajustar las influencias de variables que se utilizan para la predicción en función del contexto espacial donde se ejercen las mismas. De las dos condiciones, la de la autocorrelación espacial presentó el mayor problema en esta investigación. El efecto de ésta se redujo limitando la cantidad de casos utilizados para realizar la predicción, pasando del universo de los casos, a utilizar solo una muestra de ellos. Otro problema como el de la redundancia de variables explicables se mantenía en el modelo. La reducción de variables explicativas eliminando aquellas que resultaban ser redundantes permitió corregir esta situación. Estas modificaciones al modelo inicial permitieron que la técnica de los MCO presentara un mejor desempeño.

En cuanto al efecto del estallido de la burbuja inmobiliaria no se puede concluir que el mismo tuviera un efecto negativo en la selección de las variables, ya que se continuó identificando las mismas nueve variables como las idóneas para la predicción del valor inmueble aun pasado el momento del estallido de la burbuja.

6. LINEAS FUTURAS DE INVESTIGACION

Al igual que otras investigaciones, la presente genera nuevas preguntas lo que da paso a nuevas vías de investigación. Como se expresó en las conclusiones, una creciente cantidad de investigaciones (Brunsdon et al., 1999; Bitter et al., 2007; Yu, 2007; Ogneva-Himmelberger et al., 2009; Lu et al., 2014) han discutido la ventaja que presenta utilizar la técnica de la Regresión Geográficamente Ponderada para la predicción de valores inmuebles. Amerita tomar los datos de la presente investigación y replicar la predicción, pero utilizando la técnica de la Regresión Geográficamente Ponderada. Una vez obtenido los resultados de ésta se debe comparar los mismos con los obtenidos con la técnica de los Mínimos Cuadrados Ordinarios y determinar cual permite una mejor predicción de los valores residenciales.

7. REFERENCIAS

- Bitter, C., Mulligan, G. F., & Dall'erba, S. (2007). Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *Journal of Geographical Systems*, 9(1), 7-27.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, 39(3), 497-524

- Caridad, J. M., Núñez, J. M., & Ceular, N. (2008). Metodología de precios hedónicos vs. Redes Neuronales Artificiales como alternativas a la valoración de inmuebles. Un caso real. *CT Catastro*, 62, 27-42.
- Dorsey, R. E., Hu, H., Mayer, W. J., & Wang, H.-c. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, 19(2), 75-93.
- Du, H., & Mulley, C. (2006). Relationship Between Transport Accessibility and Land Value: Local Model Approach with Geographically Weighted Regression. *Transportation Research Record: Journal of the Transportation Research Board*, 1977(-1), 197-205.
- Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic Autocorrelation in House Prices. *The Journal of Real Estate Finance and Economics*, 23(1), 5-30.
- Kestens, Y., Thériault, M., & Des Rosiers, F. (2006). Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems*, 8(1), 61-96.
- Klovers, K., & Pereira, A. (2015). *Estimating the effect of transit on residential property values: the case of the Portland MAX system*. Retrieved from
- Krugman, P. (2009, 2 de Septiembre). How Did Economists Get It So Wrong? *New York Times*. Retrieved from <http://www.nytimes.com/2009/09/06/magazine/06Economic-t.html>
- Liu, Y., Zheng, B., Turkstra, J., & Huang, L. (2010). A hedonic model comparison for residential land value analysis. *International Journal of Applied Earth Observation and Geoinformation*, 12(Supplement 2), S181-S193.
- Ogneva-Himmelberger, Y., Pearsall, H., & Rakshit, R. (2009). Concrete evidence & geographically weighted regression: a regional analysis of wealth and the land cover in Massachusetts. *Applied Geography*, 29(4), 478-487.
- Rosenshein, L., Scott, L., & Pratt, M. (2011). Finding a Meaningful Model. *ArcUser*, Winter 2011, 5.
- Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. *Journal of Transport Geography*, 41, 53-62.
- Shiller, R. J. (2011, Febrero 5). Housing Bubbles Are Few and Far Between. *New York Times*.
- Suriatini, I. (2006). Spatial Autocorrelation and Real Estate Studies: A Literature Review. *Malaysian Journal of Real Estate*, 1(1), 1- 13.
- Tu, Y., Yu, S.-M., & Sun, H. (2004). Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach. *Real Estate Economics*, 32(2), 297-328.
- Yaolin, L., Tao, D., Yanfang, L. (2006). *A GIS-based hedonic modeling of urban land value spatio-temporal patterns* (Vol. 6366). Bellingham, WA, ETATS-UNIS: Society of Photo-Optical Instrumentation Engineers.
- Yu, D. (2007). Modeling owner-occupied single-family house values in the city of Milwaukee: A geographically weighted regression approach. *GIScience & Remote Sensing*, 44(3), 267-282.